FIGURE 1

09/463705

```
7910  CTGCTTCACGCTCACCAGCCGTATCCTTTCTCGGTTCCTCTTGTGTGCTCACTGCAACCAGG  7969
      +----+----+----+----+----+----+
      GACGAAGTGCGAGTGGTCGGCATAGGAAAGAGCCAAGGAGAACACGAGTGACGTTGGTCC
      Q  K  V  S  V                                  Q  L  W
                  <---ORF6                                      *

7970  CTTCCGGGCGCCGCCGGAGGCCACCGCGGGGGAAGATCTCGTCCAGTTCGGACAGCG       8029
      +----+----+----+----+----+----+
      GAAGGCCGCGGCGGCCTCCGGTGGCGCCCCTTCTAGAGCAGGTCAAGCCTGTCGC
      A  E  P  A  A  G  G  S  A  V  A  P  F  I  E  D  L  E  S  L

8030  CCTGCTCGTCCAGGTCATCGCGGACGCCTTCAGCGCGGAGTCGAGCTGCTCGGGGTTC      8089
      +----+----+----+----+----+----+
      GGACGAGCAGGTCCCAGTAGCGCCTGCGGAAGTCGCGCCTCAGCTCGACGAGCCCCAAG
      A  Q  E  D  L  T  M  A  S  A  K  L  A  S  D  L  Q  E  P  T

8090  GCGGGGCCGATGACGGCCCGGGCGATGCCGGGACCACAGCACCCATGCCAGCCCCACCT      8149
      +----+----+----+----+----+----+
      CGCCCCGGCTACTGCCGCGCTACGGCCCGCTGGTCGTGGGTACGCTCGGGGTGGA
      R  P  G  I  V  A  G  A  I  G  P  R  S  L  V  W  A  L  G  V

8150  CGGCCCGGGTCTTTCGCCCGAGGTTGCGGCAGAACTTCTTCGTAGGCCTCGATTCGCCGGGGCCA  8209
      +----+----+----+----+----+----+
      GCCGGGCCCAGAAGCGGTCTCCAACGCGTCTTGAAGAGCATCCGGAGCGGCCTAGCGGCCCGCGT
      E  A  P  D  E  G  L  N  R  C  F  K  E  Y  A  E  I  A  P  R

8210  GGGACGGCAACAGCACCTGGCACCGGCCCTGGCCCGCCGACTTCACCGGCGGTGCCCGCGGCCA   8269
      +----+----+----+----+----+----+
      CCCTGCCGTTGTCGTGGACCGTGGCCGGGACCGGGCGGCTGAAGTGGCCGCCACGGGCGCCGGT
      L  S  P  L  L  V  Q  A  R  G  Q  A  S  K  V  A  T  G  A  A

8270  GCTTCTCCAGCCGCTCCGCTGAGCAGGCCGTGCAGGCCGGCGGCGACCAGGGAAGACGCCGA     8329
      +----+----+----+----+----+----+
      CGAAGAGGTCGGCGAGGGACGCTCCGCGAGGCGACTCGTCCGGCCGCCGCTGGTCCCTTCTGCGGCT
      L  K  E  L  A  G  S  L  L  G  G  H  L  P  S  W  A  F  V  G
```

FIGURE 2

```
8330   GCCCGTAGGCCTGCCGGCGGCGGCAGCACCTTCCAGCTTCGGCGTGCCGGACCGCCCAGGTTGT
       +-------+-------+-------+-------+-------+-------+  8389
       CGGGCATCCGGACGCGCCGCCGCCGTCGTGGAGTTCGAGCCGCACGGCCTGGCGGTCCAACA
       L  G  Y  A  Q  A  A  P  L  V  E  L  E  A  H  R  V  A  L  N

8390   ACAGGCACTGGTGGGAGACCATGCCCAGGGAGTGGCGGGGCGTTCTCCTGCGCGG
       +-------+-------+-------+-------+-------+-------+  8449
       TGTCCGTGACCACCCTCTGGTACGGGTCCCTCACCGCCCCGCGCCAAGAGGACGCGCC
       Y  L  C  Q  H  S  V  M  G  L  S  H  R  R  A  A  N  E  Q  A

8450   CGGGCGATGTGCCAGCCCGCGAAGTTCGACGAGAGCCGACGTAGGAGACCTTGCCGCTGGCGA
       +-------+-------+-------+-------+-------+-------+  8509
       GCCGCTACACGGTCGGGCGCTTCAAGCTGCTCGGGTGCATCCTGGAACGGCGACCGCT
       A  A  I  H  W  G  A  F  N  S  S  G  V  Y  S  V  V  K  G  S  A

8510   CGAGGCTCTGTCCATGGCCTGCCCACACCTCGTCCCACGGGCGCGACCGGTCGATGTGGTGCA
       +-------+-------+-------+-------+-------+-------+  8569
       GCTCCGACAGGTACCGGACGGTGTGGAGCAGGTGCCGCGCCAGCTACACCACGT
       V  L  S  D  M  A  Q  W  V  E  D  W  P  A  S  R  D  I  H  H

8570   TCTGGTAGACGTCGATGTGGTTCGACGCCCCAGCCTTGCGCAGCGATCCCTCGCAGGAGGCGA
       +-------+-------+-------+-------+-------+-------+  8629
       AGACCATCTGCACCAGCTACACCAGCTGCCGGTCGGACGCGTCGGTAGGGAGCGTCCTCCCT
       M  Q  Y  V  D  I  H  D  V  G  L  R  R  L  S  G  E  C  S  A

8630   TGATGTGCCCGCGCCGACAGCCCGCTGTCGTTGACCGCGCTCGCTCATCTCGCCGCCGACCT
       +-------+-------+-------+-------+-------+-------+  8689
       ACTACACGGCCGCGGCCGTGTCGGGCACAGCAACTGCCGAGCGAGTAGAGCGGCGGCTGGA
       I  I  H  R  A  S  L  G  S  D  N  V  R  E  S  M  E  G  G  V

8690   TGGTCGCCAGCACGGTCCTGCGGCCGTCCGGCCGTTCCGCGCCCTGGGCCAGCCACCTGCCCACCA
       +-------+-------+-------+-------+-------+-------+  8749
       ACCAGCGGTCGTGCCACAGGAGGCGGGGACCCGGTCGGTGGACGGGTGGT
       K  T  A  L  V  T  D  E  R  R  G  G  G  G  Q  A  L  W  R  G  V
```

FIGURE 2

```
8750  GCTCCTCGGTGTGGCCCTTGTTAGAGCCGCCCAGCCGTACATGTCGGCGGTGTCGAGGCAGT  8809
      +----+----+----+----+----+----+
      CGAGGAGCCACACCGGGAACATCTCGGCGGTCGGCATGTACAGCCGCCACAGCTCCGTCA
      L  E  E  T  H  G  K  Y  L  R  W  G  Y  M  D  A  T  D  L  C

8810  TGATGCCGCGGTCCCGGGGCGTGGTCCATCAGGCGCAGCGGCGTCGTGTCGTCCTCGACGGCGTC  8869
      +----+----+----+----+----+----+
      ACTACGGCGCCAGGGCCCGACCACCAGTTAGTCCGCCGCAGCAGCAGGAGCTGCCGCAG
      N  I  G  R  D  R  A  H  D  M  L  R  L  A  D  D  D  E  V  R

8870  CGCTGAAGTTCACCGTGCCGAGCCAGAGCCTGCTGGTGAGCAGCGGGAACGCCCGAGCC  8929
      +----+----+----+----+----+----+
      GCGACTTCAAGTGGCACGGCTCGGTCTCGGACGACCACTCGTCGCGCCTTGCGGGCTCGG
      G  S  F  N  V  T  G  L  W  L  R  S  T  L  L  A  S  R  G  L

8930  GCACGTGCCGTCGCGGCGGCGTCGGTGTGGTCATCGTGGTTCTCTCCTTCCTGCGGCCAGTTCCTC  8989
      +----+----+----+----+----+----+
      CGTGCACGCCAGCGCCGCCGCCACCAGTAGCACCAAGAGAGGAGGACGCCGGTCAAGGAG
      R  V  H  T  A  A  D  T  T  M  *
         <---ORF7    *  R  P  E  R  R  G  A  A  L  E  E

8990  GCAGATGCCGACGACCTCGGCCGGTGACGGCTCCGCGAGCATGTCGTCGCGCATCCGCGC  9049
      +----+----+----+----+----+----+
      CGTCTACGGCTGCTGGAGCCGGCCACTGCCGAGGCGCTCGTACAGCAGCGCGTAGGCGCG
      C  I  G  V  V  E  A  P  S  P  E  A  L  M  D  D  R  M  R  A

9050  CGCGCCGGCGGCGGGTGGGCCCGGGTCGTCGTGGAGGACCCGCTTCACCGACTCCCGGAGCTGGTC  9109
      +----+----+----+----+----+----+
      GCGCGGCCGCCGCCCACCCGGGCCCAGCAGCTCCTGGGCGAAGTGGCTGAGGGCCTCGACCAG
      A  G  A  R  H  A  P  D  D  L  V  R  K  V  S  E  R  L  Q  D

9110  GGGGTCAGCTCGGCCACGGCAGCGGCAGCGGCGATCCCCGCCCCCGAATTCCTGCGCTGCGC  9169
      +----+----+----+----+----+----+
      CCCCAGTGAGCCGGTGCCGGTCGCGCTAGGGGCGGGGCTTAAGGACGCACGCGACCGCG
      P  T  L  E  P  V  P  L  A  I  G  A  G  F  E  Q  T  R  Q  A
```

FIGURE 2

09/463705

```
9170  GCGCACGCCGGTGTCCCAGCCGTCGGCAGGATCACCTGCGGCACGCCGTGGATCGCCGC  9229
      +----+----+----+----+----+----+
      CGCGTGCGGCCACAGGGTCGGCAGCCGTCCTAGTGGACGCCGTGCGGCACCTAGCGGCG
       R  V  G  T  D  W  G  D  P  L  I  V  Q  P  V  G  H  I  A  A

9230  GGTGTGCCAGCTCCCGGGTCCGCCGTGGTGCACCGTGCGCCGGCCAGGTCGGCAGCAGCGC  9289
      +----+----+----+----+----+----+
      CCACACGGTCGAGGGCCCAGGCGGCACCACGTGGCACGCGGCCGGTCCAGCCGTCGTCGCG
       T  H  W  S  G  P  G  G  H  V  T  A  A  C  T  P  L  L  A

9290  GTGCATCGGGACGAAGCCGACCGTGCGGACGTTGTTCCGGGATGTTCGGCACGCCTTCTAG  9349
      +----+----+----+----+----+----+
      CACGTAGCCCTGCTTCGGCTGGCAGCGCCTGCAACAGGCCCTACAAGCGCTGCGGAAGATC
       H  M  P  V  F  G  V  T  R  V  N  D  P  I  N  A  V  G  E  L

9350  CTTGCTGGCCGTCGAAGGTCGCGATGATCTCGGCCTCGACGTCCCGACGGCACCCAGCAG  9409
      +----+----+----+----+----+----+
      GACGACGCGCAGCTTCCAGCGCTACTAGAGCCGGAGCTGCAGGGCGTGCCGTGGGTCGTC
       Q  Q  A  D  F  T  A  I  I  E  A  D  V  D  G  V  A  G  L  L

9410  CTCCTCGATGGAGACCTGCCGATGCGTGTTCTCGCGGCTGGAGATCCCGAGCGTGAGGCA  9469
      +----+----+----+----+----+----+
      GAGGAGCTACCTCTGGACGGGCTACGACAAGAGCGCCGACCTCTAGGGCTCGCACTCCGT
       E  E  I  S  V  Q  G  I  S  N  E  R  S  S  I  G  L  T  L  C

9470  CACGGCCGGCGCGCTCGGCTCGTCGTCCTGCAGCCATTCCGGACGCCACCGGACGGGCCCGTTGTA  9529
      +----+----+----+----+----+----+
      GTGCCGGCCGCGCGAGCCGAGCAGCACGTCGGTAAGGCCTGCGGTGGCCTGCGGGCAACAT
       V  R  R  E  P  E  D  I  L  W  E  P  V  V  S  P  G  N  Y

9530  GTTCGACGTAGCGCCATCCCGACGGTCTTCAGGCGGTGTCGAGCCTGATCGGCCGGGGGC  9589
      +----+----+----+----+----+----+
      CAGCTGCATCGCGGTAGGGCTGCCAGAAGTCCGGCCACAGAGTCGGACTAGCCGGCCCCCG
       D  V  Y  R  M  G  V  T  K  L  G  T  D  L  R  I  A  A  P  A
```

FIGURE 2

FIGURE 2

```
10010  GATGTCGTGGCCCCGTGGGTCATGAAGTCCACGAGGTCGACGTCGGTGCCGACCGGGAC
       +-----+-----+-----+-----+-----+-----+  10069
       CTACAGCACCGGGCGCACCCAGTACTTCAGGTGCTCCAGCTGCAGCCACGGCTGGCCCTG
       I  D  H  G  A  H  T  M  F  D  V  L  D  V  D  T  G  V  P  V

10070  GGCGGTCAGCCGCCGGCCGGTGATGTCCTCGGTGAGCGCCGGGGACGGCGACCACGCGGAC
       +-----+-----+-----+-----+-----+-----+  10129
       CCGCCAGTCGGCGGCCGGCCACTACAGGAGCCACTCGCGGCCCCTGCCGTGGTGCGCCTG
       A  T  L  G  A  A  T  I  D  E  T  L  A  P  S  A  V  V  R  V

10130  CTCGTGCCCGCGCGGAACGCCCATGCGAGGGGACGAGGCCGAAGAGGTGGCTCTT
       +-----+-----+-----+-----+-----+-----+  10189
       GAGCACGGGCGGCGCCGCCTTGCGGGTACGCTCCCCCTGCTCCGGCTTCTCCACCGAGAA
       E  H  G  A  A  R  F  A  W  A  L  P  V  L  G  F  L  H  S  K

10190  GCTGGCCATGGAGGAGAAGACGACGCGCCATCGCGGTTACCTCGCGGTTACCTTCAGAGCTTCGACGGGGCAGC
       +-----+-----+-----+-----+-----+-----+  10249
       CGACCGGTACCTCCTCTTCTGCTGCGCGTAGCGCCAATGGAGTCTCGAGCTGCCCCGTCG
       S  A  M  S  S  F  V  V  R  M  *  L  E  V  P  C
                                        <---ORF8

10250  GGTTGGTTCCCCGCCAGGACGGGTGATCGGCCGGCGGCCGGCGGCCGGACGACCGGGCCGCTGGGCCGTGA
       +-----+-----+-----+-----+-----+-----+  10309
       CCAACCAAGGGGCGTCCTGCCCACTAGCCGGCCGCCGGCCGCCTGCTGGCCCGGCGACCCGCACT
       R  N  T  G  R  L  V  P  S  R  R  R  R  V  V  P  G  S  P  T

10310  GTCCGGGCAGCGCCTTGGCCCGCCCCAGTGCCGGCGGGTGGCGGGTGACCAGCT
       +-----+-----+-----+-----+-----+-----+  10369
       CAGGCCCGTCGCGGAACCGGCCGGGCGTCACGCGGCGTCACCGCTCGCCACTGGTCGA
       L  G  P  L  A  K  A  A  R  L  A  A  T  A  L  A  T  V  L

10370  CCTCCAGCCTGCCGGGGTGGCCGGCGGATGTGCCGACAGCGCCGGGTCGGGGCGGT
       +-----+-----+-----+-----+-----+-----+  10429
       GGAGGTCGGACGGCCCCACCGGCCGCCTACACGGCGTCGCGCCAGCCCCGCCCCCA
       E  E  L  R  G  P  H  G  R  H  A  S  L  A  R  D  A  D  P  R
```

FIGURE 2

```
10430  CCACGTCGAGGGCGGTCGGGCTCGGCGAAGACCTCCGGGTCGCGGTTGGCCGCCGCGACGA  10489
       +----+----+----+----+----+----+----+----+----+----+----+----+
       GGTGCAGCTCCCGCCAGCCCGAGCCGCTTCTGGAGGCCCAGCCAACCGGCGGCGCGGCTGCT
        D  V  D  L  R  D  P  E  A  F  V  E  P  D  R  N  A  A  A  V

10490  CGACCACGACCTCCTCGCCTTCGCCGATCACGTGCTCGCCGAGCCGCACCTCTGCCGGTGG  10549
       +----+----+----+----+----+----+----+----+----+----+----+----+
       GCTGGTGCTGGAGGAGCGGAAGCGGCTAGTTGCACGAGCGGCTCGGCGGTGGAGACGGCCACC
        V  V  V  V  E  E  G  E  G  I  V  H  E  G  L  R  V  E  A  T

10550  CCGTGCGCGCCGCTCCAGGTGCCAATGCCCGGGTGCAGGCGGCAGCACCTCGGGCGACGGTTCGCT  10609
       +----+----+----+----+----+----+----+----+----+----+----+----+
       GGCACGCGGCGGAGGTCCACGTTACGGGCCCAGTTCCGGCGTCGTGGAGCCGCTGCCAAGCGA
        A  T  R  R  E  L  H  L  A  P  H  L  R  L  V  E  A  V  T  R

10610  GCGCGGCGGCGGGGTCGTGCTCGGGGGCGGCGATCCGTTCGGCCAGCCCCCGGTTCGGCCGAGACGGCCA  10669
       +----+----+----+----+----+----+----+----+----+----+----+----+
       CGCGCCGCCGCCCCAGCAGCCGCTAGGCAAGCCGGTCGGGGCCAAGCCGGCTCTGCCGGT
        Q  A  A  A  P  D  D  A  I  R  E  A  L  G  P  E  A  S  V  A

10670  GGACCGCGTCGACCACGGTGTTCGCGGTCATCTCGGCGTCATCTTCGGCCCCGGCGAACAGGGCGCGCAGTG  10729
       +----+----+----+----+----+----+----+----+----+----+----+----+
       CCTGGCGCAGCTGGTGCCACAAGCGCCAGTAGAGCCGGGGGCCGCTTGTTCCCGCGCGTCAC
        L  V  A  D  V  V  T  N  A  T  M  E  A  G  A  F  L  A  R  L

10730  CGGGTCGGCGGGGCAGTGCCGCGACCGCTGCTTCGGTCACCGGAGCTGCTGGGGCTGA  10789
       +----+----+----+----+----+----+----+----+----+----+----+----+
       GCCCCAGCGCCCCGTCACGGCGCTGGGGACGAAGCCAGTGGCGCTCGACGACGCCCGACT
        A  P  D  A  P  L  A  A  V  A  A  E  T  V  A  L  Q  Q  P  S

10790  GCTGGGGCGTCCAGGCTGACGCGGGCGTCCCACGCGGGCGGCAGCACTCCGGCTGCGC  10849
       +----+----+----+----+----+----+----+----+----+----+----+----+
       CGACCCGCAGGTCCGACTGCGCCCGCAGGGTGCCGGCGCGTCGTGAGGCCGGACGCG
        L  Q  A  D  L  S  V  R  A  D  W  A  A  G  R  L  V  G  A  A
```

FIGURE 2

```
          CGAGCACGGCGGTCATGCCCTGCACCGGTACCTGCCAGGCGGAAGTCGCCGACCAGGTCCA  10909
          +----:----+----:----+----:----+----:----+----:----+----:----+
          GCTCGTGCCGCCAGTACGGGACGTGGCCATGGACGGTCCGCTTCAGCGGCTGGTCCAGGT
           G  L  V  A  T  M  G  Q  V  P  V  Q  W  A  F  D  G  V  L  D
10850

          GCCGGCGCGCCCGGCCGGGGAGCCAGACGGCGAAGCTCTCCGCCAGTTCCCCGACGTCGG  10969
          +----:----+----:----+----:----+----:----+----:----+----:----+
          CGGCCGCGCGGGCCGGCCCCTCGGTCTGCCGCTTCGAGAGGGCGGTCAAGGGGCTGCAGCC
           L  R  A  G  A  G  P  L  L  G  A  F  S  E  A  L  E  G  V  D
10910

          GGACCTCGCCTTCCCAGGACGCGGGGTGCACGTCCCCGGAACGGCTGGCCCACTCGGGCGG  11029
          +----:----+----:----+----:----+----:----+----:----+----:----+
          CCTGGAGCGGAAGGGTCCTGCGCCACGTGCAGGGCCTTGCCGACCGGGTGAGCCGCC
           P  V  E  G  E  W  S  A  A  H  V  D  R  F  P  Q  A  W  E  A
10970

          GTGGCGGCGCCCGGCGCCCGCCATCCATTCCGGTGTGTCCGGTGGCGGGGTGAACGCGG  11089
          +----:----+----:----+----:----+----:----+----:----+----:----+
          CACCGCCGCGGGGCCGGCGGTAGGTAAGGCCACACAGGCCGCCGCCACTTGCGCC
           P  P  A  G  A  A  R  M  W  E  P  T  R  G  T  A  R  T  F  A
11030

          GGTCGTCGAGCACCTGCCCGGCGCGGTGGCTGCCGTGGTCGGCCACCACCCACGTCTCGGTGCGGC  11149
          +----:----+----:----+----:----+----:----+----:----+----:----+
          CCAGCAGCTCGTGGACGGCCCCGCGCCACCGGCGGTGGTGGTGGTGCAGAGCCACGCCG
           P  D  D  L  V  Q  R  A  T  A  H  D  A  V  V  W  T  E  T  R
11090

          TGCGCCGCACACCGGACTCGCGCATCGAGCGGTTACCGGCGCTGCGGGTCGTCGTCGTGTC  11209
          +----:----+----:----+----:----+----:----+----:----+----:----+
          ACGCGGCGTGTGGCCTGAGCGCGTAGCTCGCCATGGCGCCGACGCGCCAGCAGCACACAG
           S  R  V  G  S  E  R  M  S  R  Y  R  R  Q  P  D  D  D  H
11150

          GGCGTGTCGTAGCCCATTCCCAGCGGCAACGACGGCATTGGGGTCACGTCCGGCGCCT  11269
          +----:----+----:----+----:----+----:----+----:----+----:----+
          CGCACAGCAGCATCGGGTAAGGGTCGCCGTTGCTGCCGTAACCCCAGTGCAGGCCGCGGA
           G  C  L  L  M  P  Y  P  D  G  N  S  G  Y  G  W  H  L  G  R
11210
```

FIGURE 2

```
11270    TCATCTGGAGCTGCCTGCCCAGCCCGGCGCGGATCGGTCGTGGTCATGAATTC    11321
         +----+----+----+----+----+----+----+----+----+----+----+-
         AGTAGACCTCGACGGACGGGTCGGGCCCGGCCTAGCCAGCACCAGTACTTAAG
          I  M  Q  L  Q  R  G  L  G  A  R  D  T  T  T  M
                                            <---ORF9
                                                    *  S  N
                                                    <---ORF10 . . .
```

FIGURE 2

09/463705

```
        43610              43630              43650
TTTGACAGGTCCGCCACGCGTCCCCCTACTCGACGACCACGCAATGGGCGAACAATATAG
        43670              43690              43710
GAAGGATCAAGAGGTTGACATCGCCTCGTCGAGCCAACGAACCTGTGAACATCTGCATGT
        43730              43750              43770
TGACAAGATCAACGGCGGCTACCTACTGTGGTGGCCCAGTGACGGGTTGCCGCACATCGC
        43790              43810              43830
GCTGGGGAGATTCTTTGAATTTCGCCCGTAGCACCGACCTGGAAAGCGAGCAAATGCTCC
        43850              43870              43890
GGTGAATGGGATCAGTGATTCCCCGCGTCAATTGATCACCCTTCTGGGCGCTTCCGGCTT
        V  N  G  I  S  D  S  P  R  Q  L  I  T  L  L  G  A  S  G  F
        ORF13 --->
        43910              43930              43950
CGTCGGGAGCGCGGTTCTGCGCGAGCTGCGCGACCACCCGGTCCGGCTGCGCGCGGTGTC
        V  G  S  A  V  L  R  E  L  R  D  H  P  V  R  L  R  A  V  S
        43970              43990              44010
CCGCGGCGGAGCGCCCGCGGTTCCGCCCGGCGCCGCGGAGGTCGAGGACCTGCGCGCCGA
        R  G  G  A  P  A  V  P  P  G  A  A  E  V  E  D  L  R  A  D
        44030              44050              44070
CCTGCTGGAACCGGGCCGGGCCGCCGCCGCGATCGAGGACGCCGACGTGATCGTGCACCT
        L  L  E  P  G  R  A  A  A  A  I  E  D  A  D  V  I  V  H  L
        44090              44110              44130
GGTGGCGCACGCAGCGGGCGGTTCCACCTGGCGCAGCGCCACCTCCGACCCGGAAGCCGA
        V  A  H  A  A  G  G  S  T  W  R  S  A  T  S  D  P  E  A  E
        44150              44170              44190
GCGGGTCAACGTCGGCCTGATGCACGACCTCGTCGGCGCGCTGCACGATCGCCGCAGGTC
        R  V  N  V  G  L  M  H  D  L  V  G  A  L  H  D  R  R  R  S
        44210              44230              44250
GACGCCGCCCGTGTTGCTCTACGCGAGCACCGCACAGGCCGCGAACCCGTCGGCGGCCAG
        T  P  P  V  L  L  Y  A  S  T  A  Q  A  A  N  P  S  A  A  S
        44270              44290              44310
CAGGTACGCGCAGCAGAAGACCGAGGCCGAGCGCATCCTGCGCAAAGCCACCGACGAGGG
        R  Y  A  Q  Q  K  T  E  A  E  R  I  L  R  K  A  T  D  E  G
        44330              44350              44370
CCGGGTGCGCGGCGTGATCCTGCGGCTGCCCGCCGTCTACGGCCAGAGCGGCCCGTCCGG
        R  V  R  G  V  I  L  R  L  P  A  V  Y  G  Q  S  G  P  S  G
        44390              44410              44430
CCCCATGGGGCGGGGCGTGGTCGCAGCGATGATCCGGCGTGCCCTCGCCGGCGAGCCGCT
        P  M  G  R  G  V  V  A  A  M  I  R  R  A  L  A  G  E  P  L
        44450              44470              44490
CACCATGTGGCACGACGGCGGCGTGCGCCGCGACCTGCTGCACGTCGAGGACGTGGCCAC
        T  M  W  H  D  G  G  V  R  R  D  L  L  H  V  E  D  V  A  T
        44510              44530              44550
CGCGTTCGCCGCCGCGCTGGAGCACCACGACGCGCTGGCCGGCGGCACGTGGGCGCTGGG
        A  F  A  A  A  L  E  H  H  D  A  L  A  G  G  T  W  A  L  G
        44570              44590              44610
CGCCGACCGATCCGAGCCGCTCGGCGACATCTTCCGGGCCGTCTCCGGCAGCGTCGCCCG
        A  D  R  S  E  P  L  G  D  I  F  R  A  V  S  G  S  V  A  R
        44630              44650              44670
GCAGACCGGCAGCCCCGCCGTCGACGTGGTCACCGTGCCCGCGCCCGAGCACGCCGAGGC
        Q  T  G  S  P  A  V  D  V  V  T  V  P  A  P  E  H  A  E  A
```

FIGURE 3

```
        44690                  44710                  44730
CAACGACTTCCGCAGCGACGACATCGACTCCACCGAGTTCCGCAGCCGGACCGGCTGGCG
 N  D  F  R  S  D  D  I  D  S  T  E  F  R  S  R  T  G  W  R
        44750                  44770                  44790
CCCCCGGGTTTCCCTCACCGACGGCATCGACCGGACGGTGGCCGCCCTGACCCCCACCGA
 P  R  V  S  L  T  D  G  I  D  R  T  V  A  A  L  T  P  T  E
        44810                  44830                  44850
GGAGCACTAGTGCGGGTACTGCTGACGTCCTTCGCGCACCGCACGCACTTCCAGGGACTG
 E  H  *
             V  R  V  L  L  T  S  F  A  H  R  T  H  F  Q  G  L
             ORF14 --->
        44870                  44890                  44910
GTCCCGCTGGCGTGGGCGCTGCGCACCGCGGGTCACGACGTGCGCGTGGCCGCCCAGCCC
 V  P  L  A  W  A  L  R  T  A  G  H  D  V  R  V  A  A  Q  P
        44930                  44950                  44970
GCGCTCACCGACGCGGTCATCGGCGCCGGTCTCACCGCGGTACCCGTCGGCTCCGACCAC
 A  L  T  D  A  V  I  G  A  G  L  T  A  V  P  V  G  S  D  H
        44990                  45010                  45030
CGGCTGTTCGACATCGTCCCGGAAGTCGCCGCTCAGGTGCACCGCTACTCCTTCTACCTG
 R  L  F  D  I  V  P  E  V  A  A  Q  V  H  R  Y  S  F  Y  L
        45050                  45070                  45090
GACTTCTACCACCGCGAGCAGGAGCTGCACTCGTGGGAGTTCCTGCTCGGCATGCAGGAG
 D  F  Y  H  R  E  Q  E  L  H  S  W  E  F  L  L  G  M  Q  E
        45110                  45130                  45150
GCCACCTCGCGGTGGGTATACCCGGTGGTCAACAACGACTCCTTCGTCGCCGAGCTGGTC
 A  T  S  R  W  V  Y  P  V  V  N  N  D  S  F  V  A  E  L  V
        45170                  45190                  45210
GACTTCGCCCGGGACTGGCGTCCTGACCTGGTGCTCTGGGAGCCGTTCACCTTCGCCGGC
 D  F  A  R  D  W  R  P  D  L  V  L  W  E  P  F  T  F  A  G
        45230                  45250                  45270
GCCGTCGCGGCCCGGGCCTGCGGAGCCGCGCACGCCCGGCTGCTGTGGGGCAGCGACCTC
 A  V  A  A  R  A  C  G  A  A  H  A  R  L  L  W  G  S  D  L
        45290                  45310                  45330
ACCGGCTACTTCCGCGGCCGGTTCCAGGCGCAACGCCTGCGACGGCCGCCGGAGGACCGG
 T  G  Y  F  R  G  R  F  Q  A  Q  R  L  R  R  P  P  E  D  R
        45350                  45370                  45390
CCGGACCCGCTGGGCACGTGGCTGACCGAGGTCGCGGGGCGCTTCGGCGTCGAATTCGGC
 P  D  P  L  G  T  W  L  T  E  V  A  G  R  F  G  V  E  F  G
        45410                  45430                  45450
GAGGACCTCGCCGGTCGGGCAGTGGTCGGTCGACCAGTTGCCGCCGAGTTTCCGGCTGGAC
 E  D  L  A  V  G  Q  W  S  V  D  Q  L  P  P  S  F  R  L  D
        45470                  45490                  45510
ACCGGAATGGAAACCGTTGTCGCGCGGACCCTGCCCTACAACGGCGCGTCGGTGGTTCCG
 T  G  M  E  T  V  V  A  R  T  L  P  Y  N  G  A  S  V  V  P
        45530                  45550                  45570
GACTGGCTCAAGAAGGGCAGTGCGACTCGACGCATCTGCATTACCGGAGGGTTCTCCGGA
 D  W  L  K  K  G  S  A  T  R  R  I  C  I  T  G  G  F  S  G
        45590                  45610                  45630
CTCGGGCTCGCCGCCGATGCCGATCAGTTCGCGCGGACGCTCGCGCAGCTCGCGCGATTC
 L  G  L  A  A  D  A  D  Q  F  A  R  T  L  A  Q  L  A  R  F
        45650                  45670                  45690
GATGGCGAAATCGTGGTTACGGGTTCCGGTCCGGATACCTCCGCGGTACCGGACAACATT
 D  G  E  I  V  V  T  G  S  G  P  D  T  S  A  V  P  D  N  I
```

FIGURE 3

```
        45710              45730                45750
CGTTTGGTGGATTTCGTTCCGATGGGCGTTCTGCTCCAGAACTGCGCGGCGATCATCCAC
 R  L  V  D  F  V  P  M  G  V  L  L  Q  N  C  A  A  I  I  H
        45770              45790                45810
CACGGCGGGGCCGGAACCTGGGCCACGGCACTGCACCACGGAATTCCGCAAATATCAGTT
 H  G  G  A  G  T  W  A  T  A  L  H  H  G  I  P  Q  I  S  V
        45830              45850                45870
GCACATGAATGGGATTGCATGCTACGCGGCCAGCAGACCGCGGAACTGGGCGCGGGAATC
 A  H  E  W  D  C  M  L  R  G  Q  Q  T  A  E  L  G  A  G  I
        45890              45910                45930
TACCTCCGGCCGGACGAGGTCGATGCCGACTCATTGGCGAGCGCCCTCACCCAGGTGGTC
 Y  L  R  P  D  E  V  D  A  D  S  L  A  S  A  L  T  Q  V  V
        45950              45970                45990
GAGGACCCCACCTACACCGAGAACGCGGTGAAGCTTCGCGAGGAGGCGCTGTCCGACCCG
 E  D  P  T  Y  T  E  N  A  V  K  L  R  E  E  A  L  S  D  P
        46010              46030                46050
ACGCCGCAGGAGATCGTCCCGCGACTGGAGGAACTCACGCGCCGCCACGCCGGCTAGCGG
 T  P  Q  E  I  V  P  R  L  E  E  L  T  R  R  H  A  G  *
        46070              46090                46110
TTTCCGACCGACAAGTCCGTCCGACAGCACACCTCCGGAGGGAGCAGGGATGTACGAGGG
                                                  M  Y  E  G
                                                  ORF15 --->
        46130              46150                46170
CGGGTTCGCCGAGCTTTACGACCGGTTCTACCGCGGCCGGGGCAAGGACTACGCGGCCGA
 G  F  A  E  L  Y  D  R  F  Y  R  G  R  G  K  D  Y  A  A  E
        46190              46210                46230
GGCCGCGCAGGTCGCGCGGCTGGTCAGAGACCGCCTGCCCTCGGCTTCCTCGCTGCTCGA
 A  A  Q  V  A  R  L  V  R  D  R  L  P  S  A  S  S  L  L  D
        46250              46270                46290
CGTGGCCTGCGGGACCGGCACCCACCTGCGCCGGTTCGCCGACCTCTTCGACGACGTGAC
 V  A  C  G  T  G  T  H  L  R  R  F  A  D  L  F  D  D  V  T
        46310              46330                46350
CGGGCTGGAGCTGTCGGCGGCGATGATCGAGGTCGCCCGGCCGCAGCTCGGCGGCATCCC
 G  L  E  L  S  A  A  M  I  E  V  A  R  P  Q  L  G  G  I  P
        46370              46390                46410
GGTGCTGCAGGGCGACATGCGCGACTTCGCGCTGGATCGCGAGTTCGACGCCGTCACCTG
 V  L  Q  G  D  M  R  D  F  A  L  D  R  E  F  D  A  V  T  C
        46430              46450                46470
CATGTTCAGCTCCATCGGGCACATGCGCGACGGCGCCGAGCTGGACCAGGCGCTGGCGTC
 M  F  S  S  I  G  H  M  R  D  G  A  E  L  D  Q  A  L  A  S
        46490              46510                46530
CTTCGCCCGCCACCTCGCCCCCGGCGGCGTCGTGGTGGTCGAACCGTGGTGGTTCCCGGA
 F  A  R  H  L  A  P  G  G  V  V  V  V  E  P  W  W  F  P  E
        46550              46570                46590
GGACTTCCTCGACGGCTACGTGGCCGGTGACGTGGTGCGCGACGGCGACCTGACGATCTC
 D  F  L  D  G  Y  V  A  G  D  V  V  R  D  G  D  L  T  I  S
        46610              46630                46650
GCGCGTCTCGCACTCCGTGCGCGCCGGCGGCGCGACCCGGATGGAGATCCACTGGGTCGT
 R  V  S  H  S  V  R  A  G  G  A  T  R  M  E  I  H  W  V  V
        46670              46690                46710
GGCCGACGCGGTGAACGGTCCGCGGCACCACGTGGAGCACTACGAGATCACGCTCTTCGA
 A  D  A  V  N  G  P  R  H  H  V  E  H  Y  E  I  T  L  F  E
```

FIGURE 3

09/463705

```
        46730                    46750                     46770
GCGGCAGCAGTACGAGAAGGCCTTCACCGCGGCCGGTTGCGCTGTGCAGTACCTGGAGGG
   R  Q  Q  Y  E  K  A  F  T  A  A  G  C  A  V  Q  Y  L  E  G
        46790                    46810                     46830
CGGACCCTCCGGACGCGGGTTGTTCGTCGGTGTGCGCGGATGACCCGTGCGTTCGCGTTT
   G  P  S  G  R  G  L  F  V  G  V  R  G  *
        46850                    46870                     46890
TCCGTTCCTGGCACAGGTGATCCGCTCCACGGGCCCTTTCCCCGCCGTGACCGGACCCTT
        46910                    46930                     46950
ACAGTGAGTGCGGGTCTTGATCGACAACGCCCGGCGGCAGCAAGCGGAGCCGTCGACGAC
           V  R  V  L  I  D  N  A  R  R  Q  Q  A  E  P  S  T  T
        ORF16 --->
        46970                    46990                     47010
ACCGCAGGGAGAGTCGATGGGTGATCGGACCGGCGACCGGACGATTCCGGAATCCTCGCA
   P  Q  G  E  S  M  G  D  R  T  G  D  R  T  I  P  E  S  S  Q
        47030                    47050                     47070
GACCGCAACGCGTTTCCTGCTCGGCGACGGCGGAATCCCCACCGCCACGGCGGAAACCCA
   T  A  T  R  F  L  L  G  D  G  G  I  P  T  A  T  A  E  T  H
        47090                    47110                     47130
CGACTGGCTGACCCGCAACGGCGCCGAGCAGCGGCTCGAGGTGGCGCGCGTGCCGTTCAG
   D  W  L  T  R  N  G  A  E  Q  R  L  E  V  A  R  V  P  F  S
        47150                    47170                     47190
CGCCATGGACCGCTGGTCGTTCCAGCCCGAGGACGGCAGGCTCGCCCACGAGTCCGGGCG
   A  M  D  R  W  S  F  Q  P  E  D  G  R  L  A  H  E  S  G  R
        47210                    47230                     47250
CTTCTTCTCCATCGAGGGCCTGCACGTGCGGACGAACTTCGGCTGGCGGCGGGACTGGAT
   F  F  S  I  E  G  L  H  V  R  T  N  F  G  W  R  R  D  W  I
        47270                    47290                     47310
CCAGCCCATCATCGTGCAGCCCGAGATCGGCTTCCTCGGCCTCATCGTCAAGGAGTTCGA
   Q  P  I  I  V  Q  P  E  I  G  F  L  G  L  I  V  K  E  F  D
        47330                    47350                     47370
CGGTGTGCTGCACGTGCTGGCGCAGGCCAAGGCCGAGCCGGGCAACATCAACGCCGTCCA
   G  V  L  H  V  L  A  Q  A  K  A  E  P  G  N  I  N  A  V  Q
        47390                    47410                     47430
GCTCTCCCCGACCCTGCAGGCGACCCGCAGCAACTACACCGGCGTCCACCGCGGCTCGAA
   L  S  P  T  L  Q  A  T  R  S  N  Y  T  G  V  H  R  G  S  K
        47450                    47470                     47490
GGTCCGGTTCATCGAGTACTTCAACGGCACGCGCCCGAGCCGGATCCTCGTCGACGTGCT
   V  R  F  I  E  Y  F  N  G  T  R  P  S  R  I  L  V  D  V  L
        47510                    47530                     47550
CCAGTCCGAGCAGGGCGCGTGGTTCCTGCGCAAGCGCAACCGGAACATGGTCGTCGAGGT
   Q  S  E  Q  G  A  W  F  L  R  K  R  N  R  N  M  V  V  E  V
        47570                    47590                     47610
GTTCGACGACCTGCCCGAGCACCCGAACTTCCGGTGGCTGACCGTCGCGCAGCTGCGGGC
   F  D  D  L  P  E  H  P  N  F  R  W  L  T  V  A  Q  L  R  A
        47630                    47650                     47670
GATGCTGCACCACGACAACGTGGTGAACATGGACCTGCGCACCGTGCTGGCCTGCGTCCC
   M  L  H  H  D  N  V  V  N  M  D  L  R  T  V  L  A  C  V  P
        47690                    47710                     47730
GACCGCCGTGGAGCGGGACCGGGCCGACGACGTGCTCGCGCGCCTGCCCGAGGGCTCGTT
   T  A  V  E  R  D  R  A  D  D  V  L  A  R  L  P  E  G  S  F
```

FIGURE 3

```
         47750                47770                47790
CCAGGCCCGGCTGCTGCACTCGTTCATCGGCGCGGGCACCCCGGCCAACAACATGAACAG
  Q  A  R  L  L  H  S  F  I  G  A  G  T  P  A  N  N  M  N  S
         47810                47830                47850
CCTGCTGAGCTGGATCTCCGACGTGCGCGCCAGGCGCGAGTTCGTGCAGCGCGGCCGCCC
  L  L  S  W  I  S  D  V  R  A  R  R  E  F  V  Q  R  G  R  P
         47870                47890                47910
GCTGCCCGACATCGAGCGCAGCGGGTGGATCCGCCGCGACGACGGCATCGAGCACGAGGA
  L  P  D  I  E  R  S  G  W  I  R  R  D  D  G  I  E  H  E  E
         47930                47950                47970
GAAGAAGTACTTCGACGTCTTCGGCGTCACGGTGGCGACCAGCGACCGCGAGGTCAACTC
  K  K  Y  F  D  V  F  G  V  T  V  A  T  S  D  R  E  V  N  S
         47990                48010                48030
GTGGATGCAGCCGCTGCTCTCGCCCGCCAACAACGGCCTGCTCGCCCTGCTGGTCAAGGA
  W  M  Q  P  L  L  S  P  A  N  N  G  L  L  A  L  L  V  K  D
         48050                48070                48090
CATCGGCGGCACGTTGCACGCGCTCGTGCAGCTGCGCACCGAGGCGGGCGGGATGGACGT
  I  G  G  T  L  H  A  L  V  Q  L  R  T  E  A  G  G  M  D  V
         48110                48130                48150
CGCCGAGCTGGCGCCTACGGTGCACTGCCAGCCCGACAACTACGCCGACGCGCCCGAGGA
  A  E  L  A  P  T  V  H  C  Q  P  D  N  Y  A  D  A  P  E  E
         48170                48190                48210
GTTCCGACCGGCCTATGTGGACTACGTGTTGAACGTGCCGCGCTCGCAGGTCCGCTACGA
  F  R  P  A  Y  V  D  Y  V  L  N  V  P  R  S  Q  V  R  Y  D
         48230                48250                48270
CGCATGGCACTCCGAGGAGGGCGGCCGGTTCTACCGCAACGAGAACCGGTACATGCTGAT
  A  W  H  S  E  E  G  G  R  F  Y  R  N  E  N  R  Y  M  L  I
         48290                48310                48330
CGAGGTGCCCGCCGACTTCGACGCCAGTGCCGCTCCCGACCACCGGTGGATGACCTTCGA
  E  V  P  A  D  F  D  A  S  A  A  P  D  H  R  W  M  T  F  D
         48350                48370                48390
CCAGATCACCTACCTGCTCGGGCACAGCCACTACGTCAACATCCAGCTGCGCAGCATCAT
  Q  I  T  Y  L  L  G  H  S  H  Y  V  N  I  Q  L  R  S  I  I
         48410                48430                48450
CGCGTGCGCCTCGGCCGTCTACACCAGGACCGCCGGATGAAACGCGCTGACCGACCTG
  A  C  A  S  A  V  Y  T  R  T  A  G  *
                                        M  K  R  A  L  T  D  L
                                        ORF17  --->
         48470                48490                48510
GCGATCTTCGGCGGCCCCGAGGCATTCCTGCACACCCTCTACGTGGGCAGGCCGACCGTC
  A  I  F  G  G  P  E  A  F  L  H  T  L  Y  V  G  R  P  T  V
         48530                48550                48570
GGGGACCGGGAGCGGTTCTTCGCCCGCCTGGAGTGGGCGCTGAACAACAACTGGCTGACC
  G  D  R  E  R  F  F  A  R  L  E  W  A  L  N  N  N  W  L  T
         48590                48610                48630
AACGGCGGACCACTGGTGCGCGAGTTCGAGGGCCGGGTCGCCGACCTGGCGGGTGTCCGC
  N  G  G  P  L  V  R  E  F  E  G  R  V  A  D  L  A  G  V  R
         48650                48670                48690
CACTGCGTGGCCACCTGCAACGCGACGGTCGCGCTGCAACTGGTGCTGCGCGCGAGCGAC
  H  C  V  A  T  C  N  A  T  V  A  L  Q  L  V  L  R  A  S  D
         48710                48730                48750
GTGTCCGGCGAGGTCGTCATGCCTTCGATGACGTTCGCGGCCACCGCGCACGCGGCGAGC
  V  S  G  E  V  V  M  P  S  M  T  F  A  A  T  A  H  A  A  S
```

FIGURE 3

```
        48770                48790                     48810
TGGCTGGGGCTGGAACCGGTGTTCTGCGACGTGGACCCCGAGACCGGCCTGCTCGACCCC
 W  L  G  L  E  P  V  F  C  D  V  D  P  E  T  G  L  L ·D  P
        48830                48850                     48870
GAGCACGTCGCGTCGCTGGTGACACCGCGGACGGGCGCGATCATCGGCGTGCACCTGTGG
 E  H  V  A  S  L  V  T  P  R  T  G  A  I  I  G  V  H  L  W
        48890                48910                     48930
GGCAGGCCCGCTCCGGTCGAGGCGCTGGAGAAGATCGCCGCCGAGCACCAGGTCAAACTC
 G  R  P  A  P  V  E  A  L  E  K  I  A  A  E  H  Q  V  K  L
        48950                48970                     48990
TTCTTCGACGCCGCGCACGCGCTGGGCTGCACCGCCGGCGGGCGGCCGGTCGGCGCCTTC
 F  F  D  A  A  H  A  L  G  C  T  A  G  G  R  P  V  G  A  F
        49010                49030                     49050
GGCAACGCCGAGGTGTTCAGCTTCCACGCCACGAAGGCGGTCACCTCGTTCGAGGGCGGC
 G  N  A  E  V  F  S  F  H  A  T  K  A  V  T  S  F  E  G  G
        49070                49090                     49110
GCCATCGTCACCGACGACGGGCTGCTGGCCGACCGCATCCGCGCCATGCACAACTTCGGG
 A  I  V  T  D  D  G  L  L  A  D  R  I  R  A  M  H  N  F  G
        49130                49150                     49170
ATCGCACCGGACAAGCTGGTGACCGATGTCGGCACCAACGGCAAGATGAGCGAGTGCGCC
 I  A  P  D  K  L  V  T  D  V  G  T  N  G  K  M  S  E  C  A
        49190                49210                     49230
GCGGCGATGGGCCTCACCTCGCTCGACGCCTTCGCCGAGACCAGGGTGCACAACCGCCTC
 A  A  M  G  L  T  S  L  D  A  F  A  E  T  R  V  H  N  R  L
        49250                49270                     49290
AACCACGCGCTCTACTCCGACGAGCTCCGCGACGTGCGCGGCATATCCGTGCACGCGTTC
 N  H  A  L  Y  S  D  E  L  R  D  V  R  G  I  S  V  H  A  F
        49310                49330                     49350
GATCCTGGCGAGCAGAACAACTACCAGTACGTGATCATCTCGGTGGACTCCGCGGCCACC
 D  P  G  E  Q  N  N  Y  Q  Y  V  I  I  S  V  D  S  A  A  T
        49370                49390                     49410
GGCATCGACCGCGACCAGTTGCAGGCGATCCTGCGAGCGGAGAAGGTTGTGGCACAACCC
 G  I  D  R  D  Q  L  Q  A  I  L  R  A  E  K  V  V  A  Q  P
        49430                49450                     49470
TACTTCTCCCCCGGGTGCCACCAGATGCAGCCGTACCGGACCGAGCCGCCGCTGCGGCTG
 Y  F  S  P  G  C  H  Q  M  Q  P  Y  R  T  E  P  P  L  R  L
        49490                49510                     49530
GAGAACACCGAACAGCTCTCCGACCGGGTGCTCGCGCTGCCCACCGGCCCCGCGGTGTCC
 E  N  T  E  Q  L  S  D  R  V  L  A  L  P  T  G  P  A  V  S
        49550                49570                     49590
AGCGAGGACATCCGGCGGGTGTGCGACATCATCCGGCTCGCCGCCACCAGCGGCGAGCTG
 S  E  D  I  R  R  V  C  D  I  I  R  L  A  A  T  S  G  E  L
        49610                49630                     49650
ATCAACGCGCAATGGGACCAGAGGACGCGCAACGGTTCGTGACGACCTGCGCCACAAGTG
 I  N  A  Q  W  D  Q  R  T  R  N  G  S  *
        49670                49690                     49710
CCAGGAGGTTCGCTCCCCGATGAACACAACTCGTACGGCAACCGCCCAGGAAGCGGGGGT
                      M  N  T  T  R  T  A  T  A  Q  E  A  G  V
                      ORF18 --->
        49730                49750                     49770
CGCCGACGCGGCGCGCCCGGACGTCGACCGGCGGGCGGTCGTGCGGGCGCTGAGCTCGGA
 A  D  A  A  R  P  D  V  D  R  R  A  V  V  R  A  L  S  S  E
```

FIGURE 3

WO 99/05283 ● ● PCT/FR98/01593

```
        49790                  49810                  49830
GGTCTCCCGCGTCACCGGCGCCGGTGACGGTGACGCCGACGTGCAGGCCGCCCGGCTCGC
  V  S  R  V  T  G  A  G  D  G  D  A  D  V  Q  A  A  R  L  A
        49850                  49870                  49890
CGACCTCGCCGCGCACTACGGGGCGCACCCGTTCACGCCGCTGGAGCAGACGCGTGCGCG
  D  L  A  A  H  Y  G  A  H  P  F  T  P  L  E  Q  T  R  A  R
        49910                  49930                  49950
GCTCGGCCTGGACCGCGCGGAGTTCGCCCACCTGCTCGACCTGTTCGGCCGCATCCCGGA
  L  G  L  D  R  A  E  F  A  H  L  L  D  L  F  G  R  I  P  D
        49970                  49990                  50010
CCTGGGCACCGCGGTGGAGCACGGTCCGGCGGGCAAGTACTGGTCCAACACGATCAAGCC
  L  G  T  A  V  E  H  G  P  A  G  K  Y  W  S  N  T  I  K  P
        50030                  50050                  50070
GCTGGACGCCGCAGGCGCACTGGACGCGGCGGTCTACCGCAAGCCTGCCTTCCCCTACAG
  L  D  A  A  G  A  L  D  A  A  V  Y  R  K  P  A  F  P  Y  S
        50090                  50110                  50130
CGTCGGCCTGTACCCCGGGCCGACGTGCATGTTCCGCTGCCACTTCTGCGTGCGGGTGAC
  V  G  L  Y  P  G  P  T  C  M  F  R  C  H  F  C  V  R  V  T
        50150                  50170                  50190
CGGTGCCCGCTACGAGGCCGCATCGGTCCCGGCGGGCAACGAGACGCTGGCCGCGATCAT
  G  A  R  Y  E  A  A  S  V  P  A  G  N  E  T  L  A  A  I  I
        50210                  50230                  50250
CGACGAGGTGCCCACGGACAACCCGAAGGCGATGTACATGTCGGGCGGGCTCGAGCCGCT
  D  E  V  P  T  D  N  P  K  A  M  Y  M  S  G  G  L  E  P  L
        50270                  50290                  50310
GACCAACCCCGGTCTCGGCGAGCTGGTGTCGCACGCCGCCGGGCGCGGTTTCGACCTCAC
  T  N  P  G  L  G  E  L  V  S  H  A  A  G  R  G  F  D  L  T
        50330                  50350                  50370
CGTCTACACCAACGCCTTCGCCCTCACCGAGCAGACGCTGAACCGCCAGCCCGGCCTGTG
  V  Y  T  N  A  F  A  L  T  E  Q  T  L  N  R  Q  P  G  L  W
        50390                  50410                  50430
GGAGCTGGGCGCGATCCGCACGTCCCTCTACGGGCTGAACAACGACGAGTACGAGACGAC
  E  L  G  A  I  R  T  S  L  Y  G  L  N  N  D  E  Y  E  T  T
        50450                  50470                  50490
CACCGGCAAGCGCGGCGCTTTCGAACGCGTCAAGAAGAACCTGCAGGGCTTCCTGCGGAT
  T  G  K  R  G  A  F  E  R  V  K  K  N  L  Q  G  F  L  R  M
        50510                  50530                  50550
GCGCGCCGAGCGGGACGCGCCGATCCGGCTCGGCTTCAACCACATCATCCTGCCGGGACG
  R  A  E  R  D  A  P  I  R  L  G  F  N  H  I  I  L  P  G  R
        50570                  50590                  50610
GGCCGACCGGCTCACCGACCTCGTCGACTTCATCGCCGAGCTCAACGAGTCCAGCCCGCA
  A  D  R  L  T  D  L  V  D  F  I  A  E  L  N  E  S  S  P  Q
        50630                  50650                  50670
ACGGCCGCTGGACTTCGTGACGGTGCGCGAGGACTACAGCGGCCGCGACGACGGCCGGCT
  R  P  L  D  F  V  T  V  R  E  D  Y  S  G  R  D  D  G  R  L
        50690                  50710                  50730
GTCGGACTCCGAGCGCAACGAGCTGCGCGAGGGCCTGGTGCGGTTCGTCGACTACGCCGC
  S  D  S  E  R  N  E  L  R  E  G  L  V  R  F  V  D  Y  A  A
        50750                  50770                  50790
CGAGCGGACCCCGGGCATGCACATCGACCTGGGCTACGCCCTGGAGAGCCTGCGGCGGGG
  E  R  T  P  G  M  H  I  D  L  G  Y  A  L  E  S  L  R  R  G
```

FIGURE 3

```
              50810                 50830                    50850
TGTGGACGCCGAGCTGCTGCGCATCCGGCCGGAGACGATGCGTCCCACCGCGCACCCCCA
  V  D  A  E  L  L  R  I  R  P  E  T  M  R  P  T  A  H  P  Q
              50870                 50890                    50910
GGTCGCGGTGCAGATCGACCTGCTCGGCGACGTCTACCTCTACCGCGAGGCGGGCTTCCC
  V  A  V  Q  I  D  L  L  G  D  V  Y  L  Y  R  E  A  G  F  P
              50930                 50950                    50970
GGAGCTGGAGGGCGCCACCCGCTACATCGCGGGCCGGGTCACCCCGTCGACCAGCCTGCG
  E  L  E  G  A  T  R  Y  I  A  G  R  V  T  P  S  T  S  L  R
              50990                 51010                    51030
CGAGGTGGTGGAGAACTTCGTGCTGGAGAACGAGGGCGTGCAGCCCCGCCCCGGCGACGA
  E  V  V  E  N  F  V  L  E  N  E  G  V  Q  P  R  P  G  D  E
              51050                 51070                    51090
GTACTTCCTCGACGGCTTCGACCAGTCGGTGACCGCACGGCTCAACCAGCTCGAACGAGA
  Y  F  L  D  G  F  D  Q  S  V  T  A  R  L  N  Q  L  E  R  D
              51110                 51130                    51150
CATCGCCGACGGGTGGGAGGACCACCGCGGCTTCCTGCGCGGAAGGTGAACCGGAGTTGC
  I  A  D  G  W  E  D  H  R  G  F  L  R  G  R  *
              51170                 51190                    51210
GAGTACGTGAGCTGGCGGTGGCGGGCGGTTTCGAGTTCACCCCCGACCCGAAGCAGGACC
                    V  A  G  G  F  E  F  T  P  D  P  K  Q  D  R
                 ORF19 --->
              51230                 51250                    51270
GGCGGGGCCTGTTCGTGTCTCCGCTGCAGGACGAGGCGTTCGTGGGCGCGGTGGGCCATC
  R  G  L  F  V  S  P  L  Q  D  E  A  F  V  G  A  V  G  H  R
              51290                 51310                    51330
GGTTCCCCGTCGCCCAGATGAACCACATCGTCTCCGCCCGGGGCGTGCTGCGCGGGCTGC
  F  P  V  A  Q  M  N  H  I  V  S  A  R  G  V  L  R  G  L  H
              51350                 51370                    51390
ACTTCACCACCACCCCGCCGGGGCAGTGCAAGTACGTCTACTGCGCGCGCGGCCGGGCGC
  F  T  T  T  P  P  G  Q  C  K  Y  V  Y  C  A  R  G  R  A  L
              51410                 51430                    51450
TCGACGTCATCGTCGACATCCGGGTCGGCTCGCCGACGTTCGGGAAGTGGGACGCGGTGG
  D  V  I  V  D  I  R  V  G  S  P  T  F  G  K  W  D  A  V  E
              51470                 51490                    51510
AGATGGACACCGAGCACTTCCGGGCGGTCTACTTCCCCAGGGGCACCGCGCACGCCTTCC
  M  D  T  E  H  F  R  A  V  Y  F  P  R  G  T  A  H  A  F  L
              51530                 51550                    51570
TCGCGCTTGAGGACGACACCCTGATGTCGTACCTGGTCAGCACGCCGTACGTGGCCGAGT
  A  L  E  D  D  T  L  M  S  Y  L  V  S  T  P  Y  V  A  E  Y
              51590                 51610                    51630
ACGAGCAGGCGATCGACCCGTTCGACCCCGCGCTGGGTCTGCCGTGGCCCGCGGACCTGG
  E  Q  A  I  D  P  F  D  P  A  L  G  L  P  W  P  A  D  L  E
              51650                 51670                    51690
AGGTCGTGCTCTCCGACCGCGACACGGTGGCCGTGGACCTGGAGACCGCCAGGCGGCGAG
  V  V  L  S  D  R  D  T  V  A  V  D  L  E  T  A  R  R  R  G
              51710                 51730                    51750
GGATGCTGCCCGACTACGCCGACTGCCTCGGCGAGGAGCCCGCCAGCACCGGCAGGTGAC
  M  L  P  D  Y  A  D  C  L  G  E  E  P  A  S  T  G  R  *
              *  A  S  Q  R  P  S  S  G  A  L  V  P  L  H
                                               <--- ORF20...
```
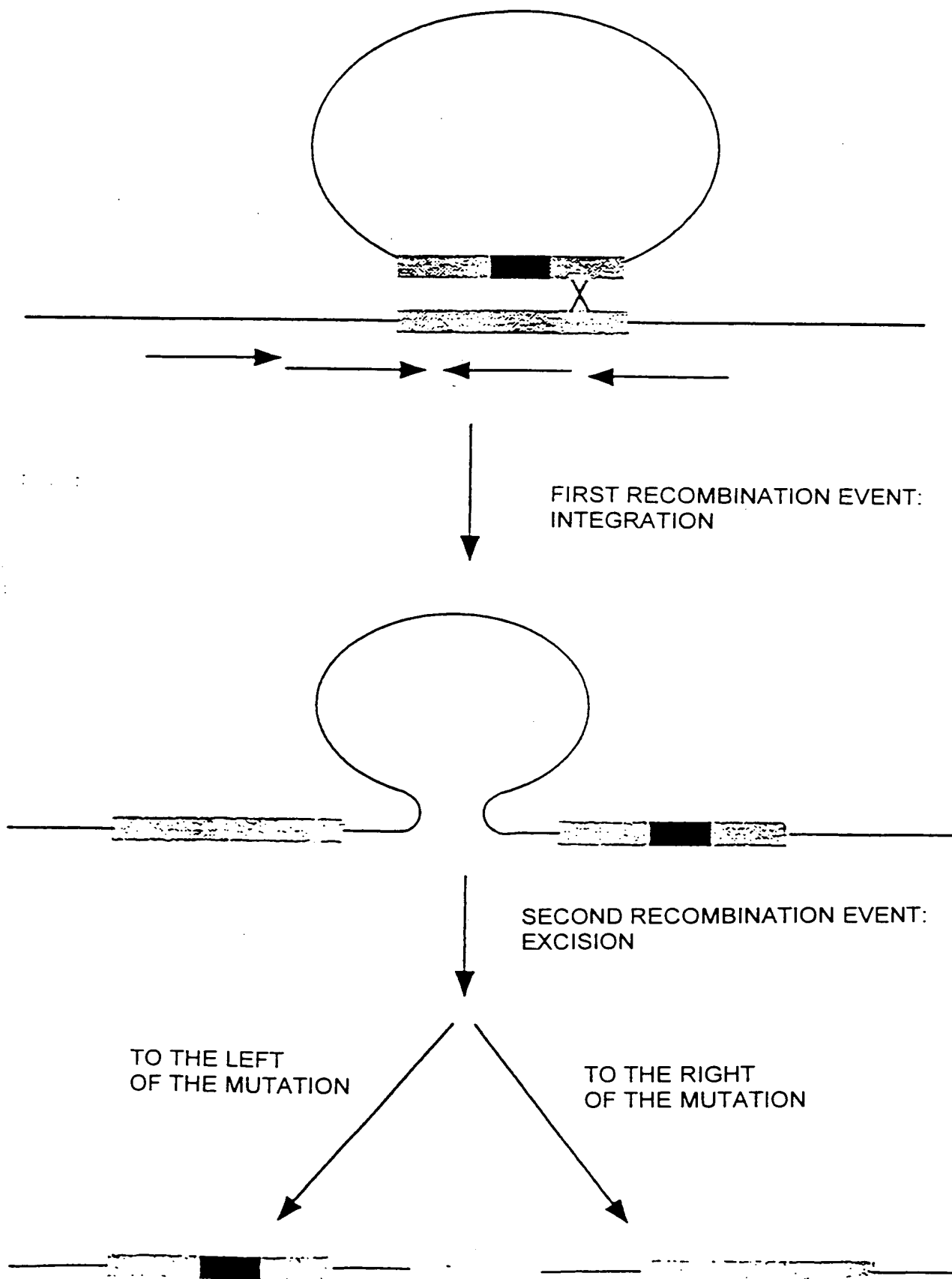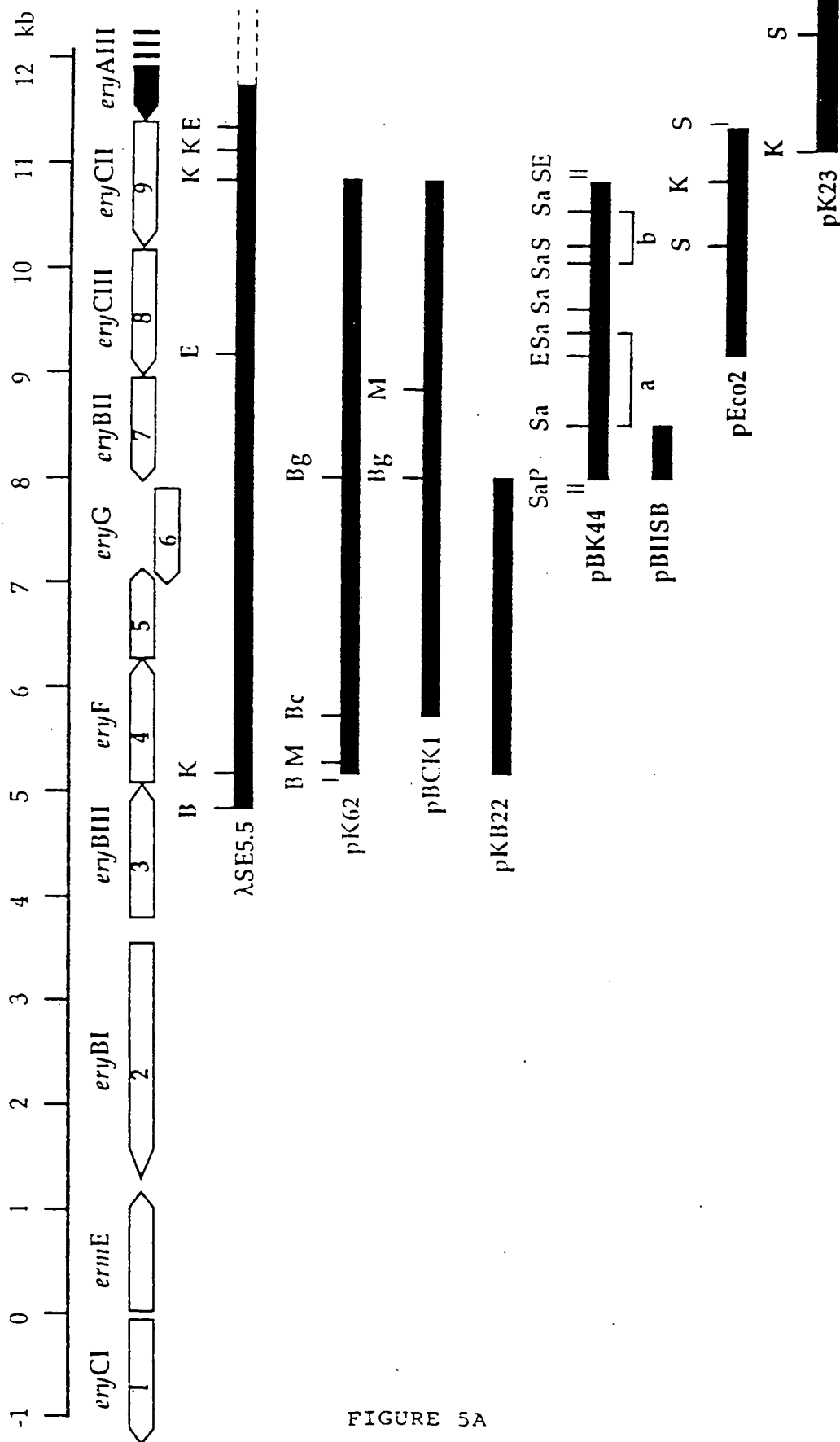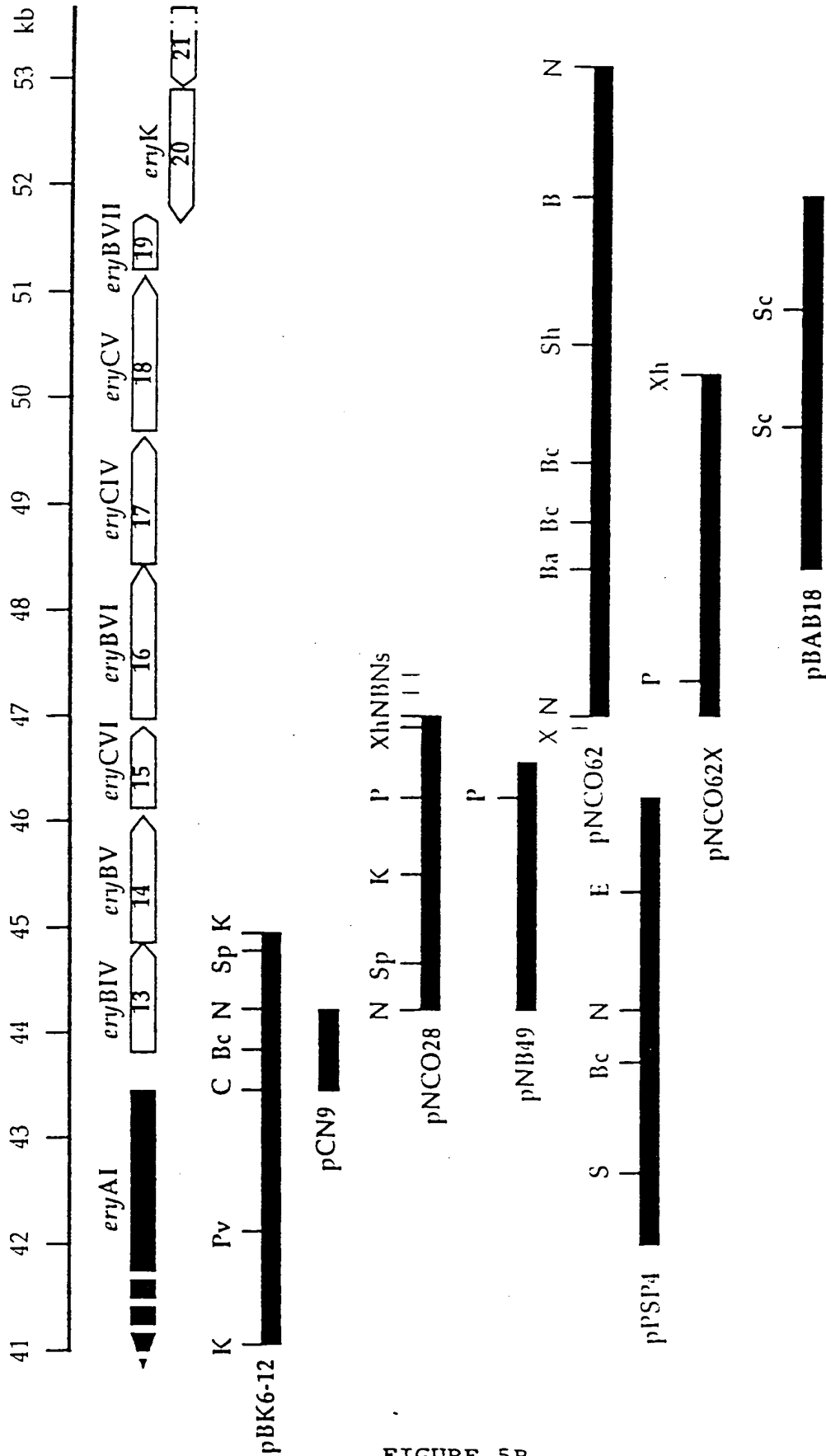
FIGURE 3

FIRST RECOMBINATION EVENT:
INTEGRATION

SECOND RECOMBINATION EVENT:
EXCISION

TO THE LEFT
OF THE MUTATION

TO THE RIGHT
OF THE MUTATION
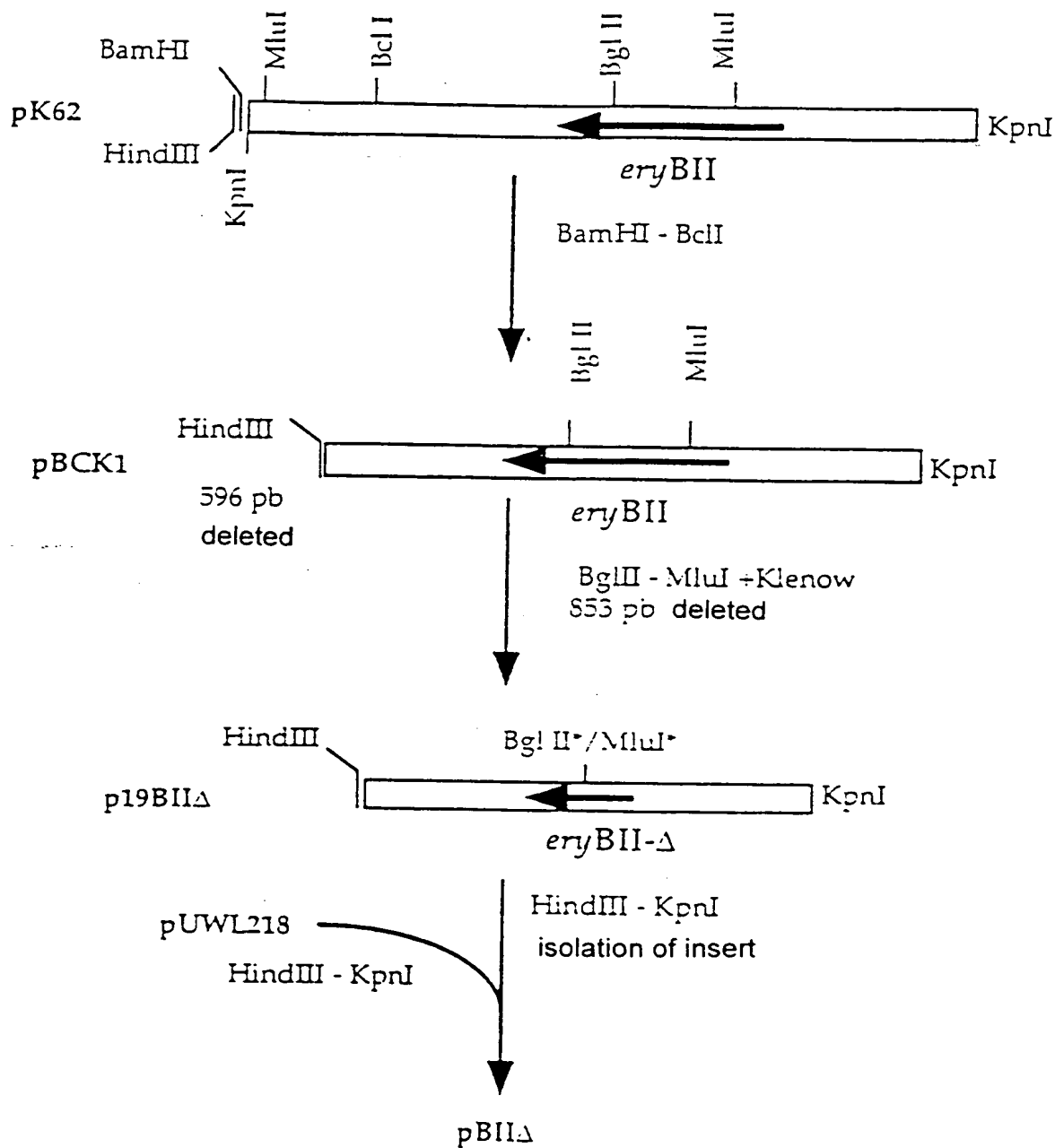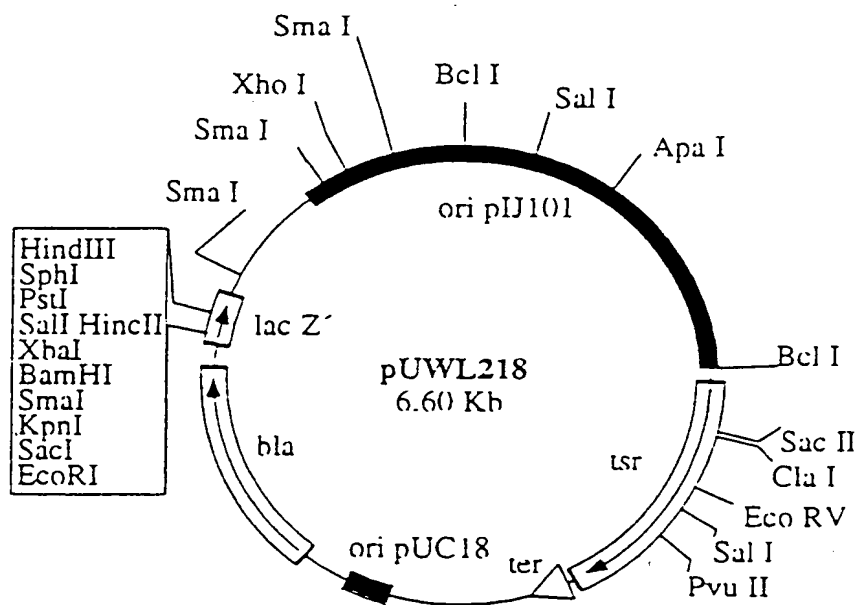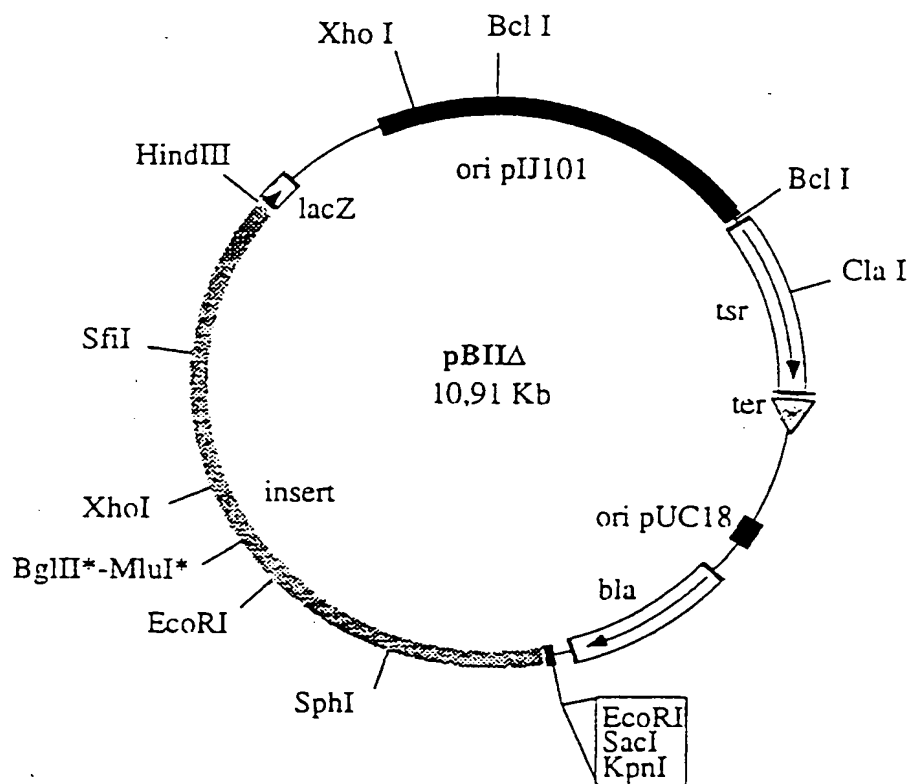
FIGURE 4
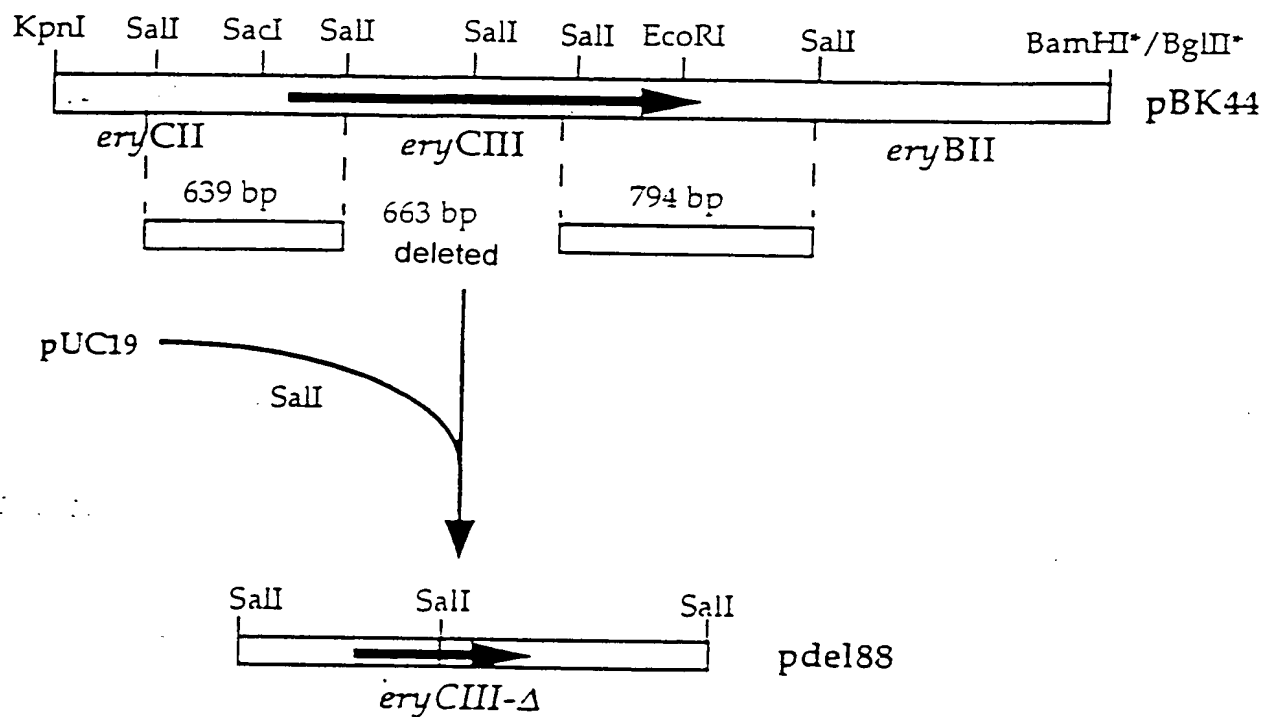
FIGURE 5A
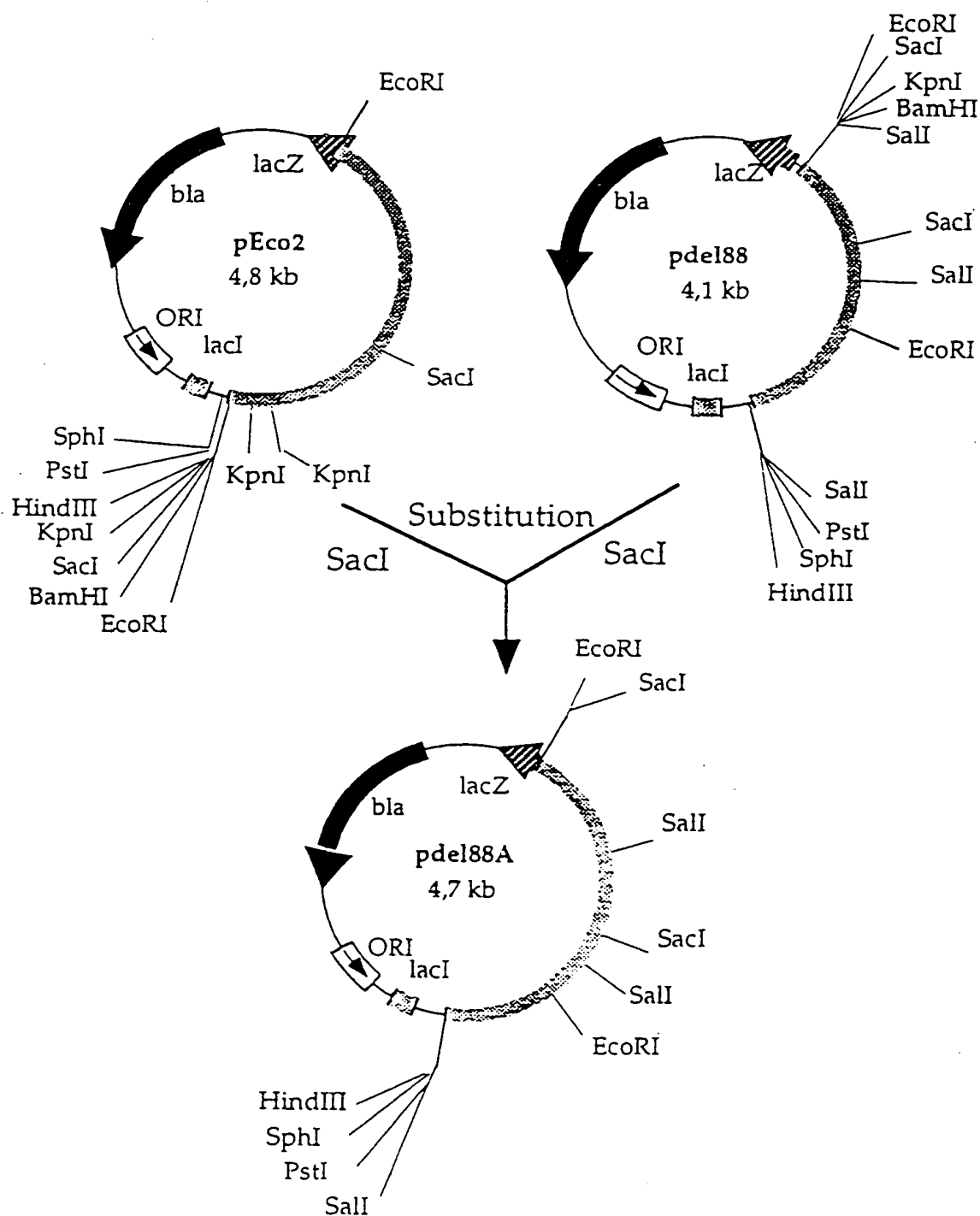
FIGURE 5B

FIGURE 6A

FIGURE 6B
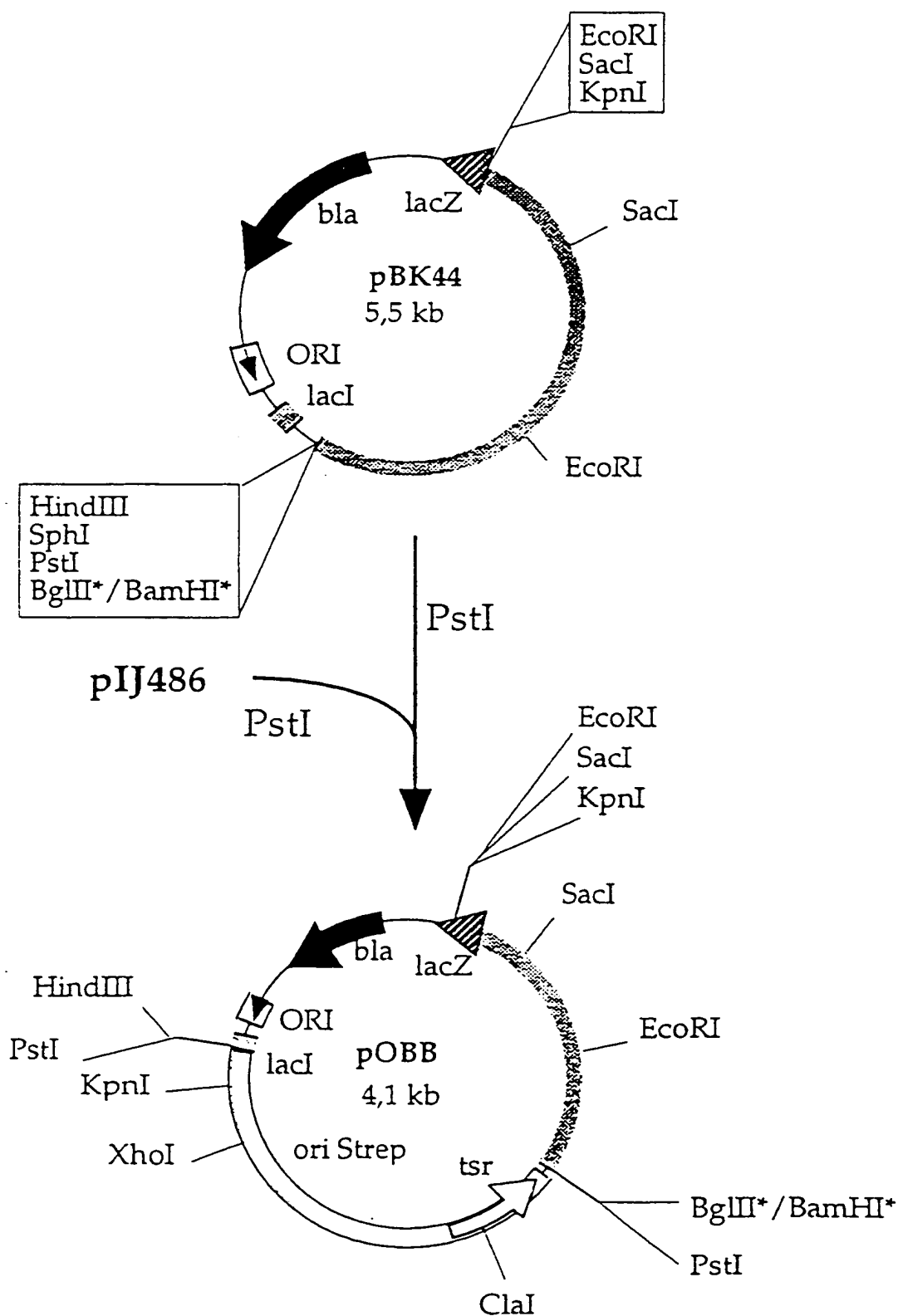
FIGURE 6C

FIGURE 7A

FIGURE 7B

FIGURE 7C

FIGURE 7D

FIGURE 8A

FIGURE 8B

FIGURE 8C
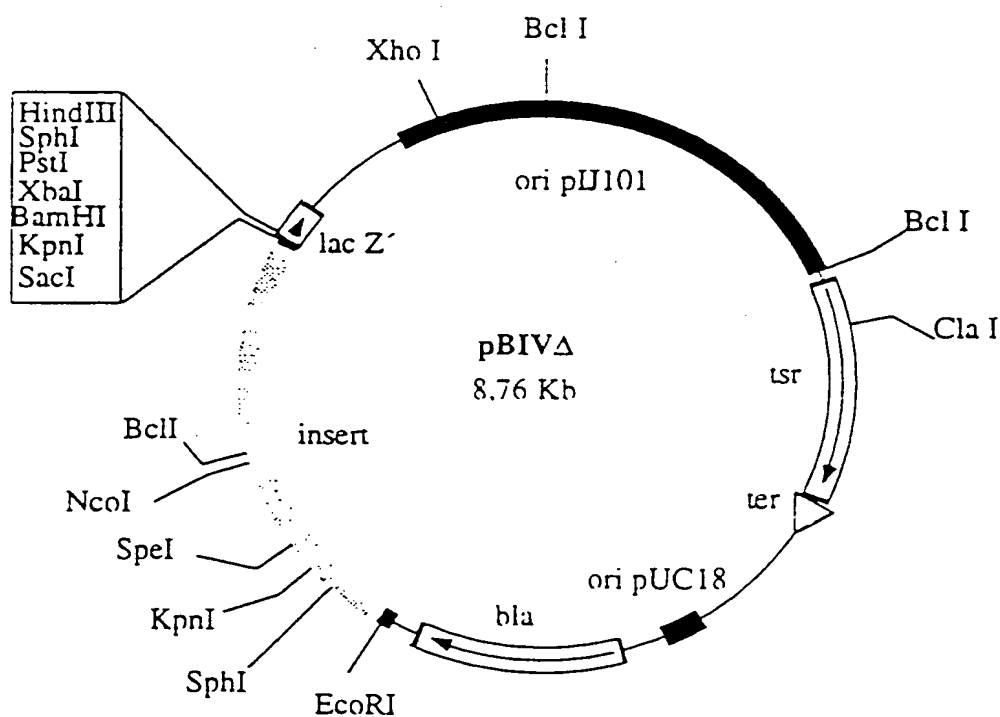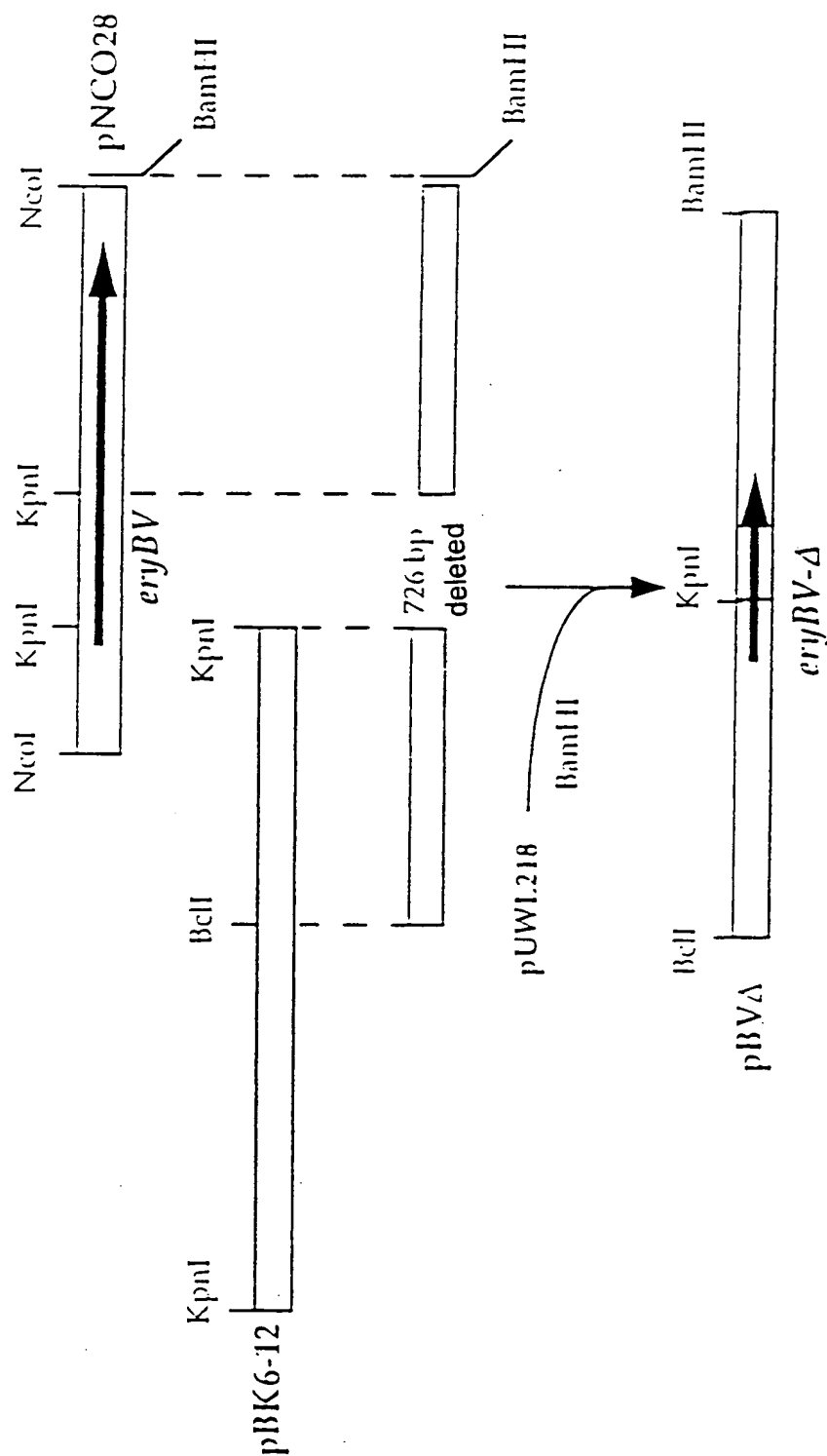
FIGURE 9A

FIGURE 9B

FIGURE 10A
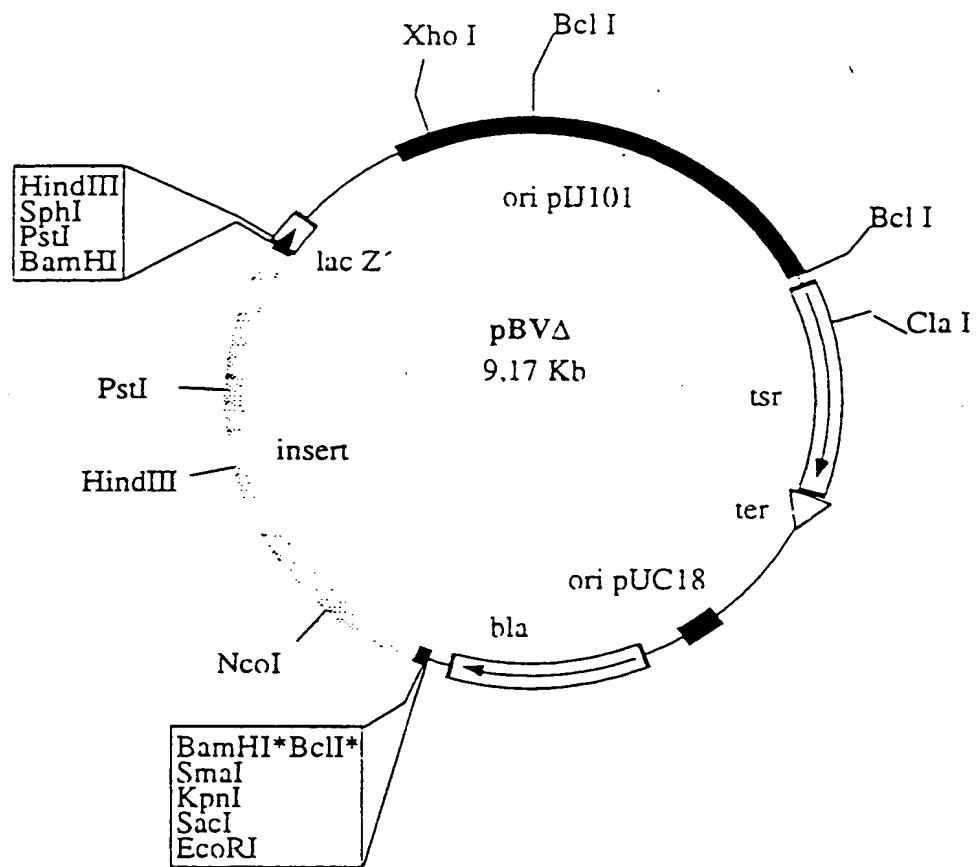
FIGURE 10B
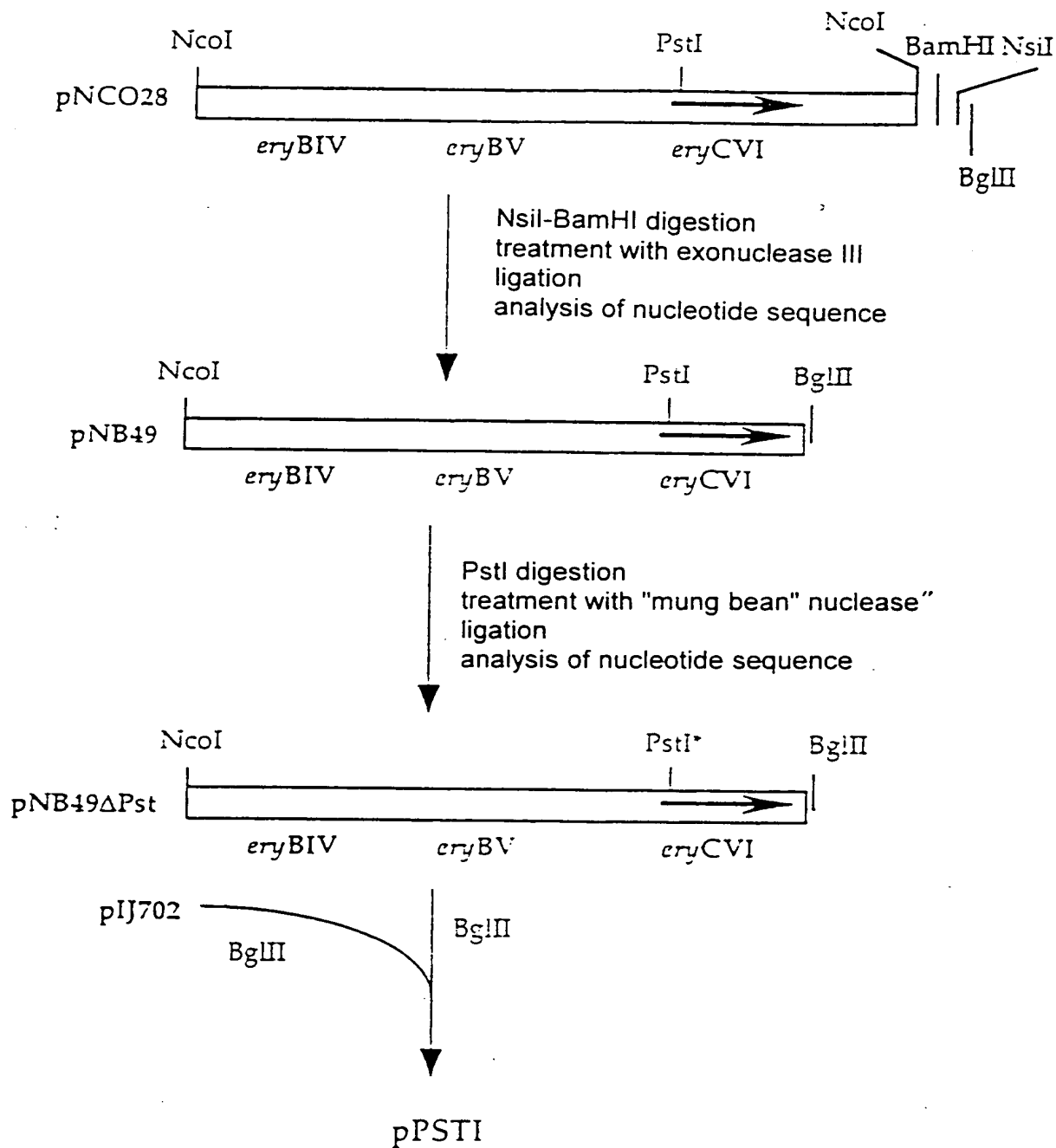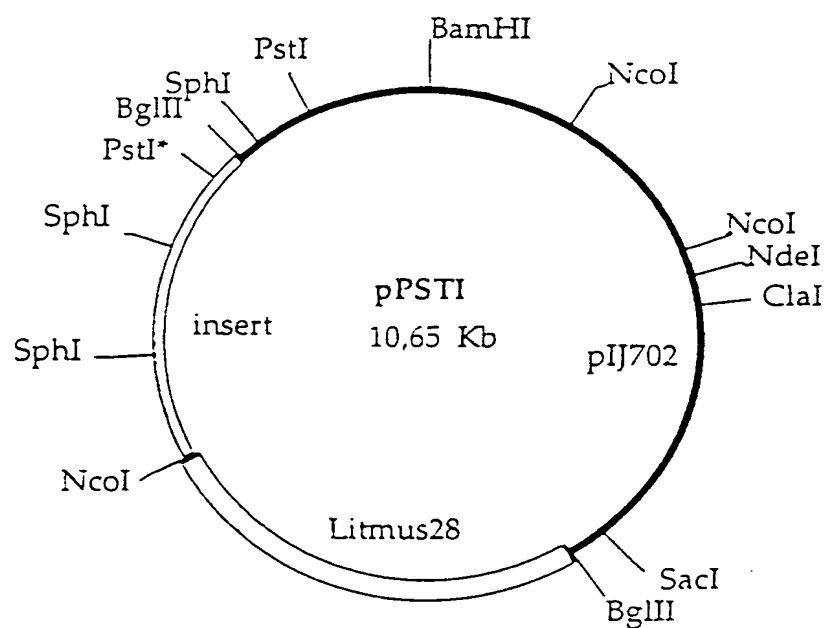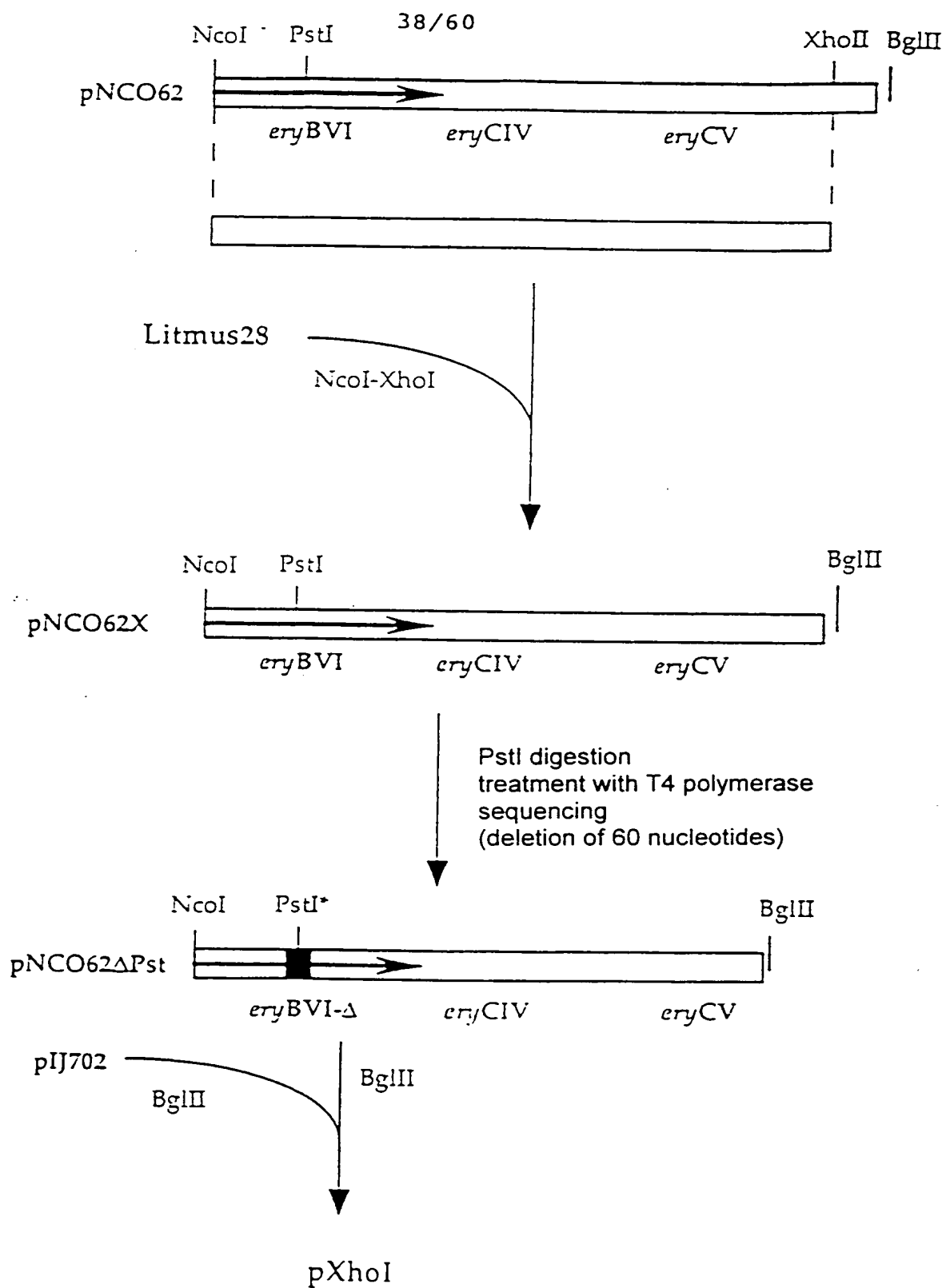
FIGURE 11A

FIGURE 11B

FIGURE 12A
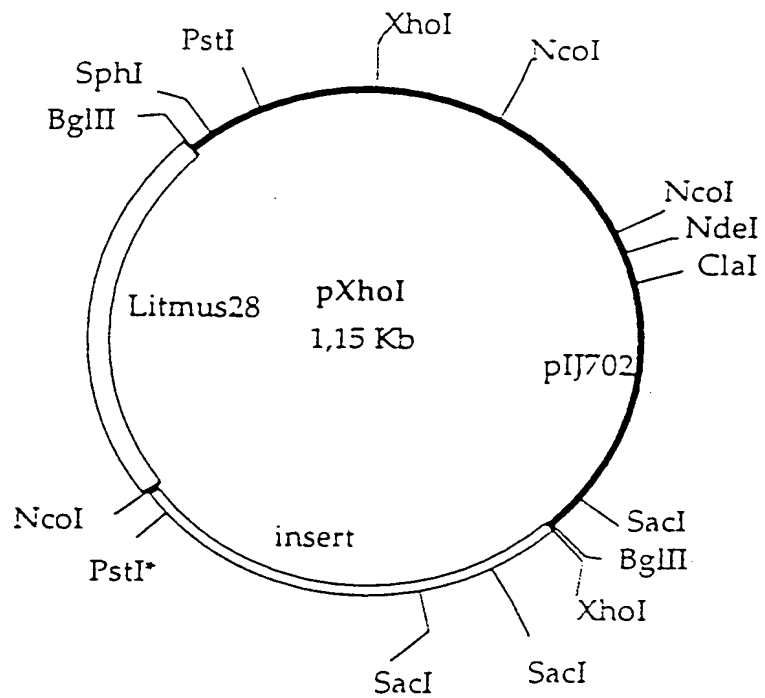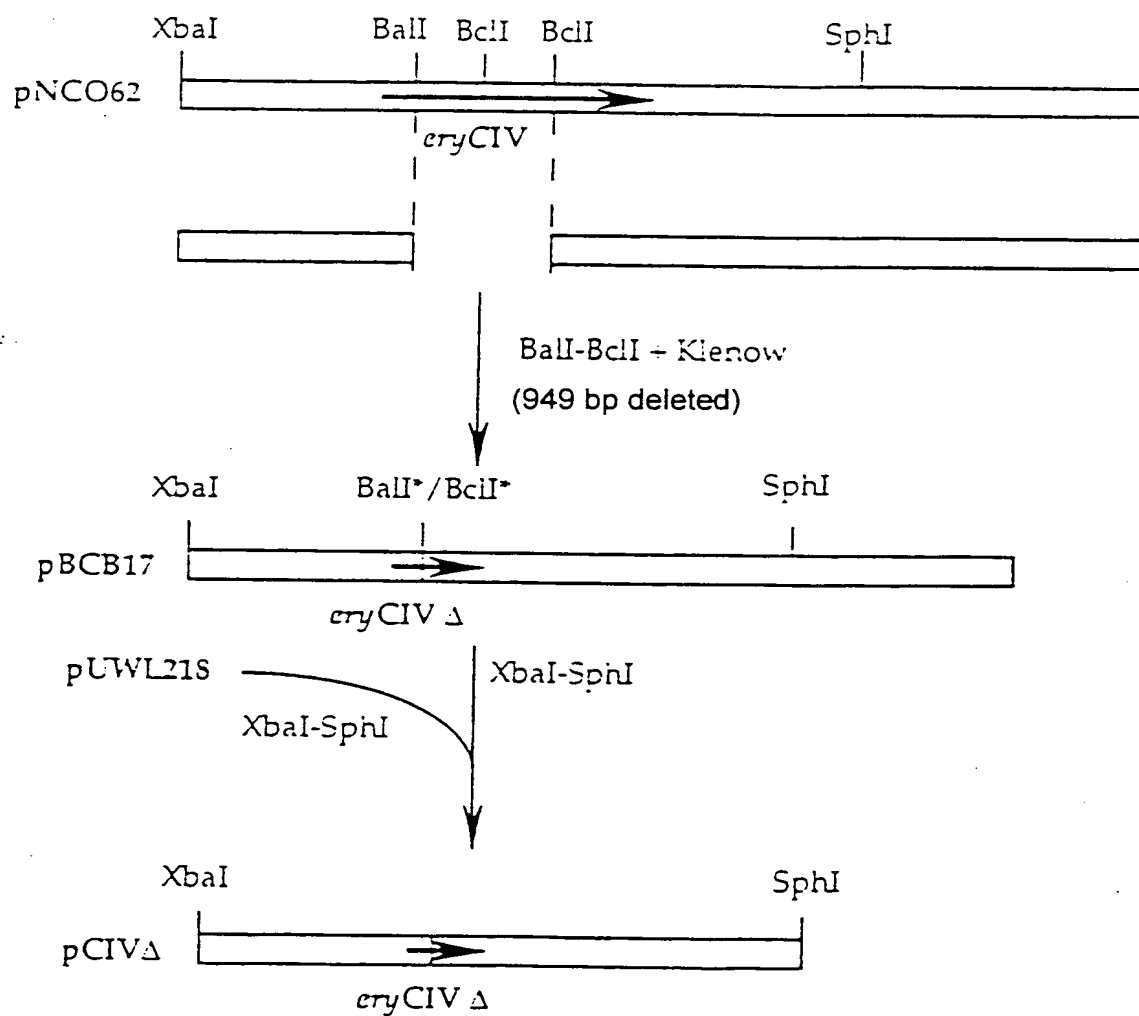
FIGURE 12B
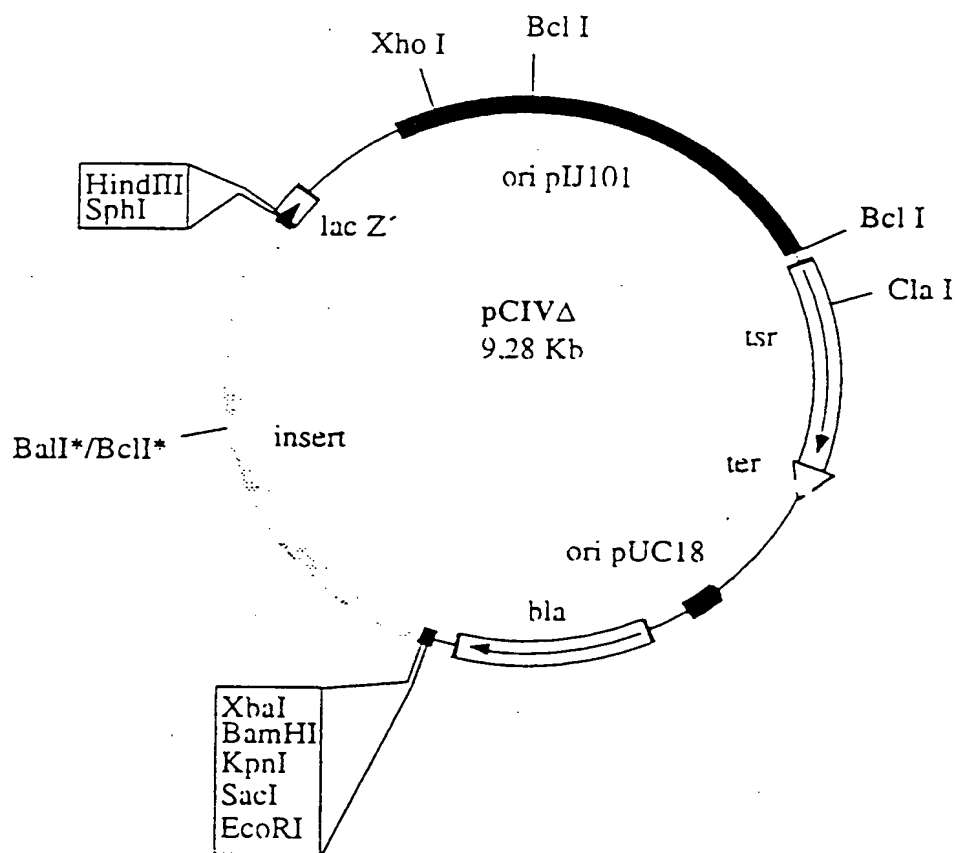
FIGURE 13A

FIGURE 13B
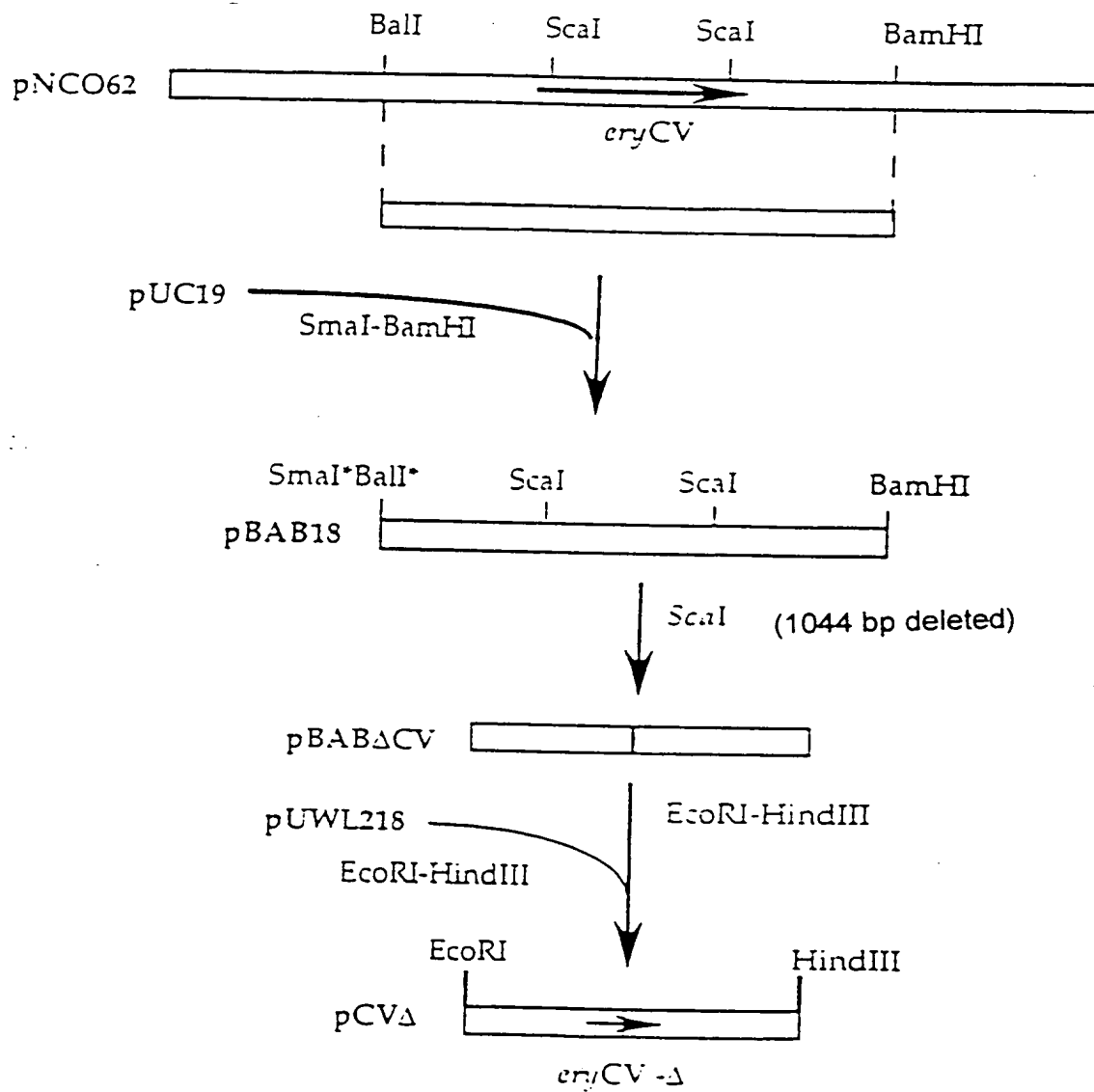
FIGURE 14A

FIGURE 14B

FIGURE 15

09/463705

FIGURE 17

FIGURE 18

1   2   3   4
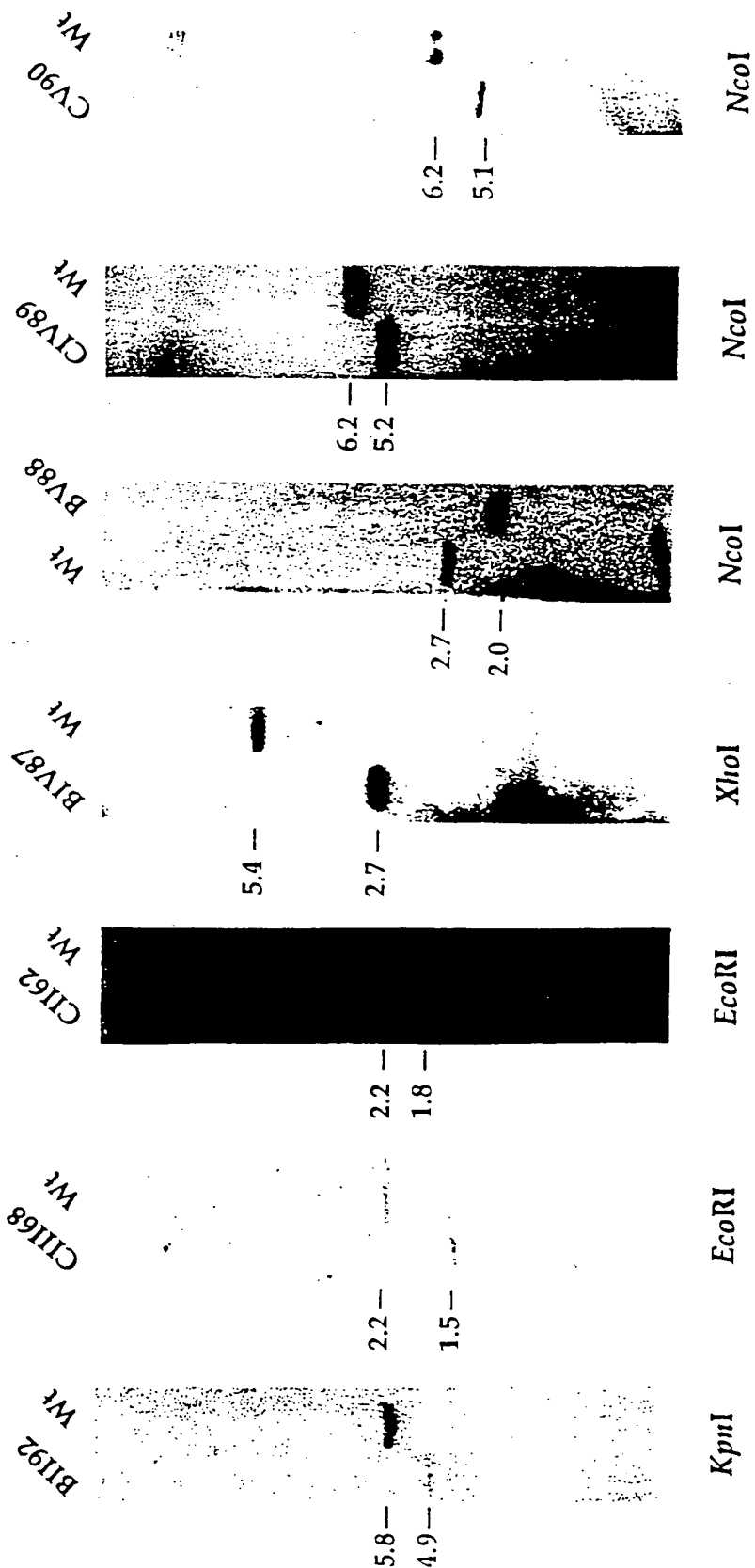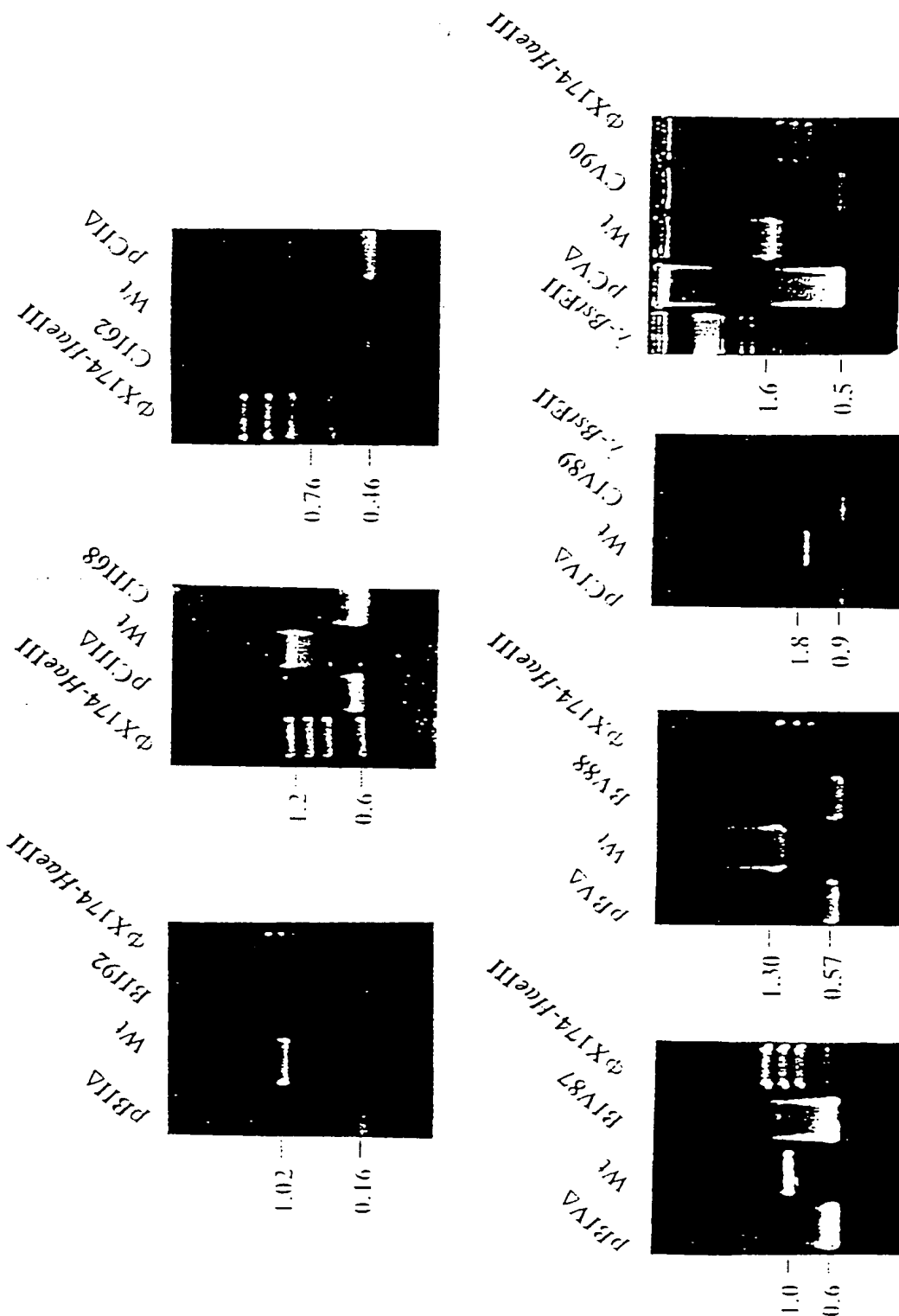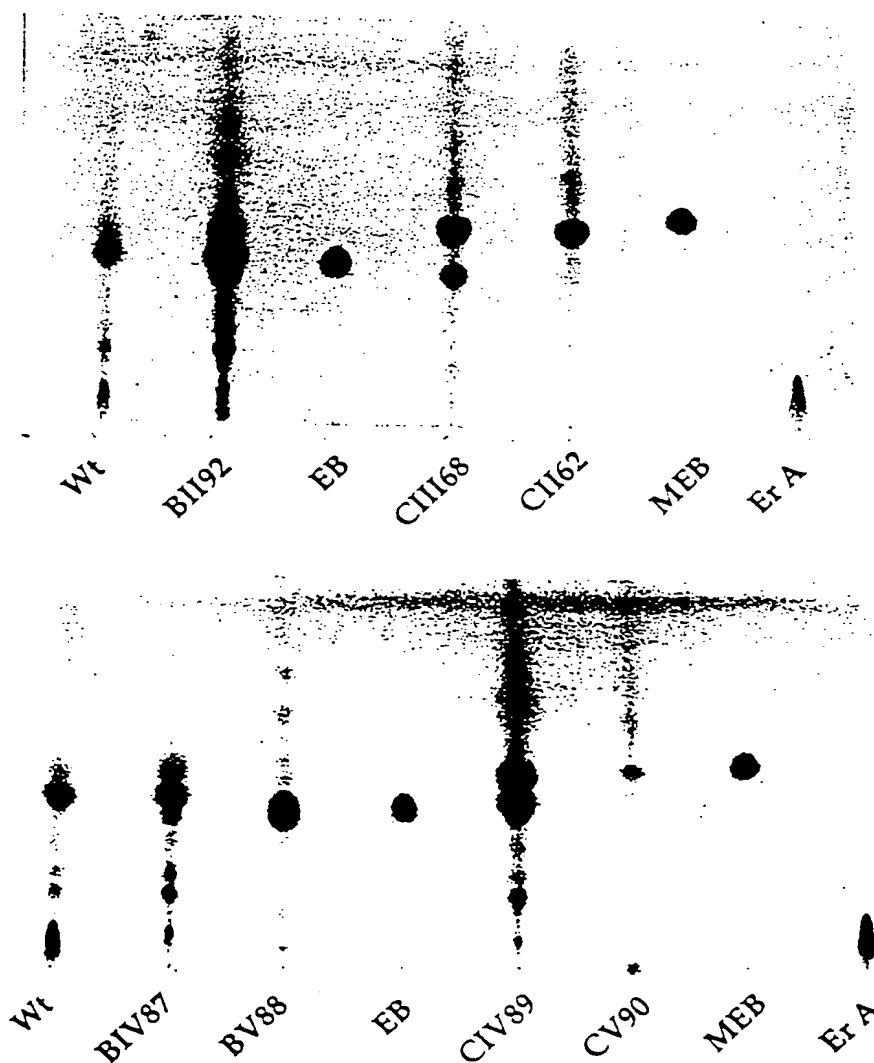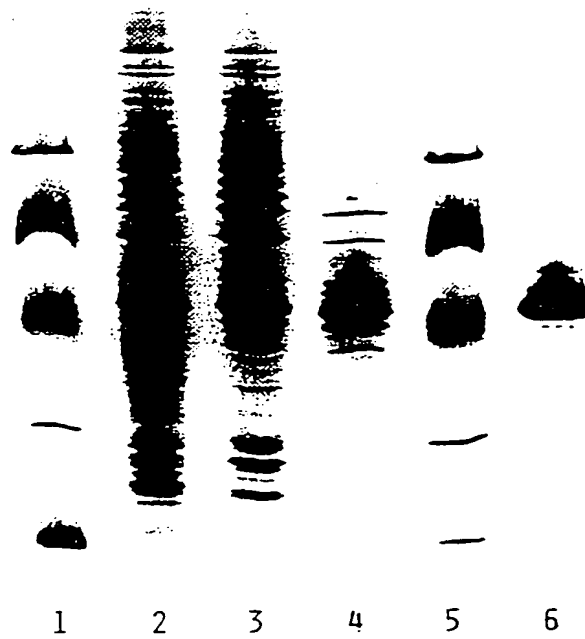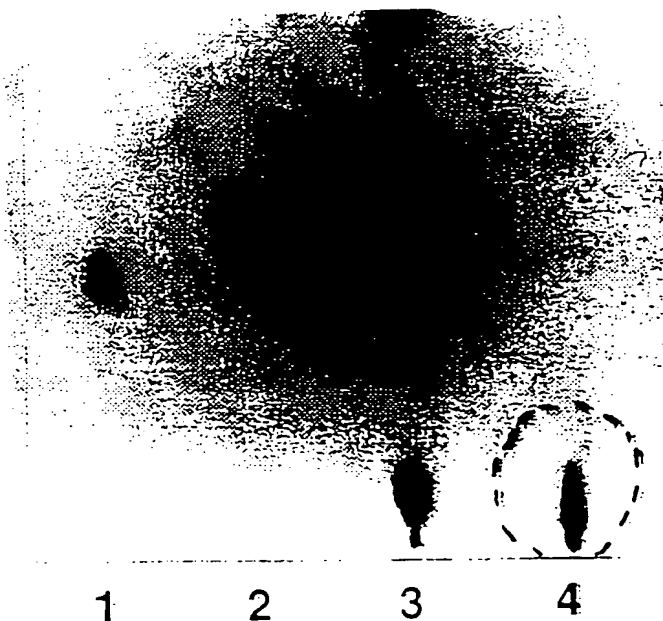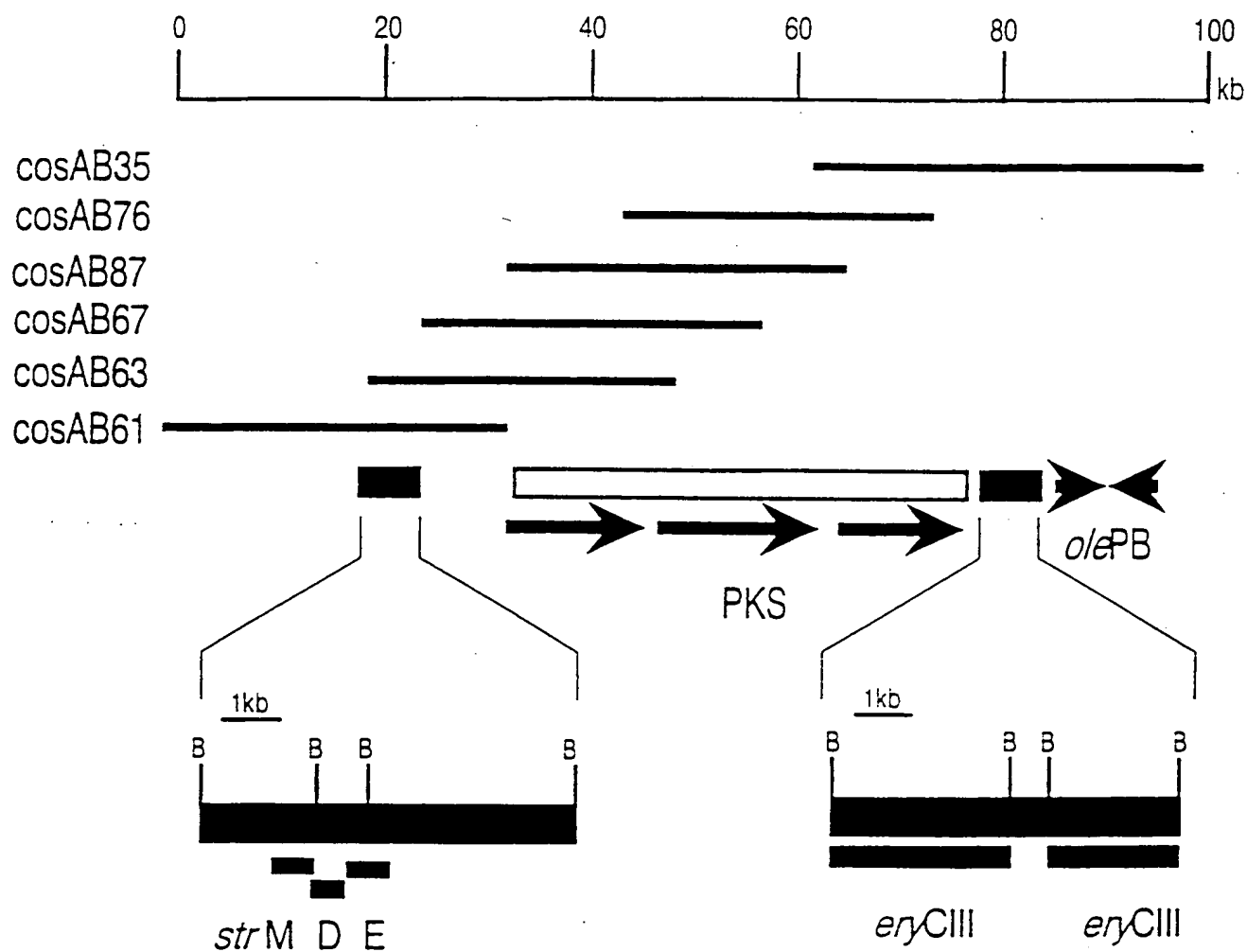
FIGURE 13

FIGURE 20

FIGURE 21

00/463705

```
1    GCATGCCCCGCTTTCCTCCCCTCTCCGAACGCCATCGACGACCCGATCCCCCTCAGGGAC   60

61   CGGTGAAGGAGCGTGTTGCACTCATGCAGGACATGCAAGGCGTACAGCCCGAACCAGCCA   120

121  GTGTCGAACACGCGGCGACGCAGCTCGAACAGAGCGAACGGCGCACGAAGCCGCCCCAG    180
                              PstI
181  GAGATGGAGGACAGCGAACTGGGGCGGCCGCCTGCAGATGCTCCGCGGCATGCAGTGGGTC  240
     M   E   D   S   E   L   G   R   R   L   Q   M   L   R   G   M   Q   W   V
     OleP1

241  TTCGGCGCCAACGGGCGATCCGTACGCCCGGCTGCTGTGTGGCATGGAGGATGACCCGTCA  300
     F   G   A   N   G   D   P   Y   A   R   L   L   C   G   M   E   D   D   P   S

301  CCTTTCTACGACGCCGATACGGCACCCTGGGCGAGCTGCACCGGAGCCGGAGCCTGG      360
     P   F   Y   D   A   I   R   T   L   G   E   L   H   R   S   R   T   G   A   W

361  GTCACCGCCGACCCCGGCGCTCGGGGGGCCGCCATCCTCGGCCGGAAGGCTCGGTGCCCG   420
     V   T   A   D   P   G   L   G   G   R   I   L   A   D   R   K   A   R   C   P

421  GAAGGCTCGTGGCCCGTGCGGGCCGAAGACCGGCAGCCAGTACGTGCTGCCCGGG       480
     E   G   S   W   P   V   R   A   K   T   D   G   L   E   Q   Y   V   L   P   G

481  CACCAGGCGTTCCTGCGGCTGGAGCGGGAGGAGGCCGAGCGACTGCGGGAGGTCGCGGCG   540
     H   Q   A   F   L   R   L   E   R   E   E   A   E   R   L   R   E   V   A   A

541  CCGGTGCTGGGGGCGGCCGCCGCGGTCGACGCGGTGGCGCTGATCGACGAGGTCTGCGCG   600
     P   V   L   G   A   A   A   V   D   A   W   R   P   L   I   D   E   V   C   A

601  GGGCTCGCCGAAGGGGCTGCCGAAGCGTTGCCACGTTCGACCTGGTCGAGGAGTACGCGGGGCTGGTG  660
     G   L   A   K   G   L   P   D   T   F   D   L   V   E   E   Y   A   G   L   V

661  CCGGTCGAGGTGCTGGCCCGGGATCTGGGGGCGTCCCGGAGGAGGACCGCGCCCGGTTCGGG  720
     P   V   E   V   L   A   R   I   W   G   V   P   E   E   D   R   A   R   F   G
```
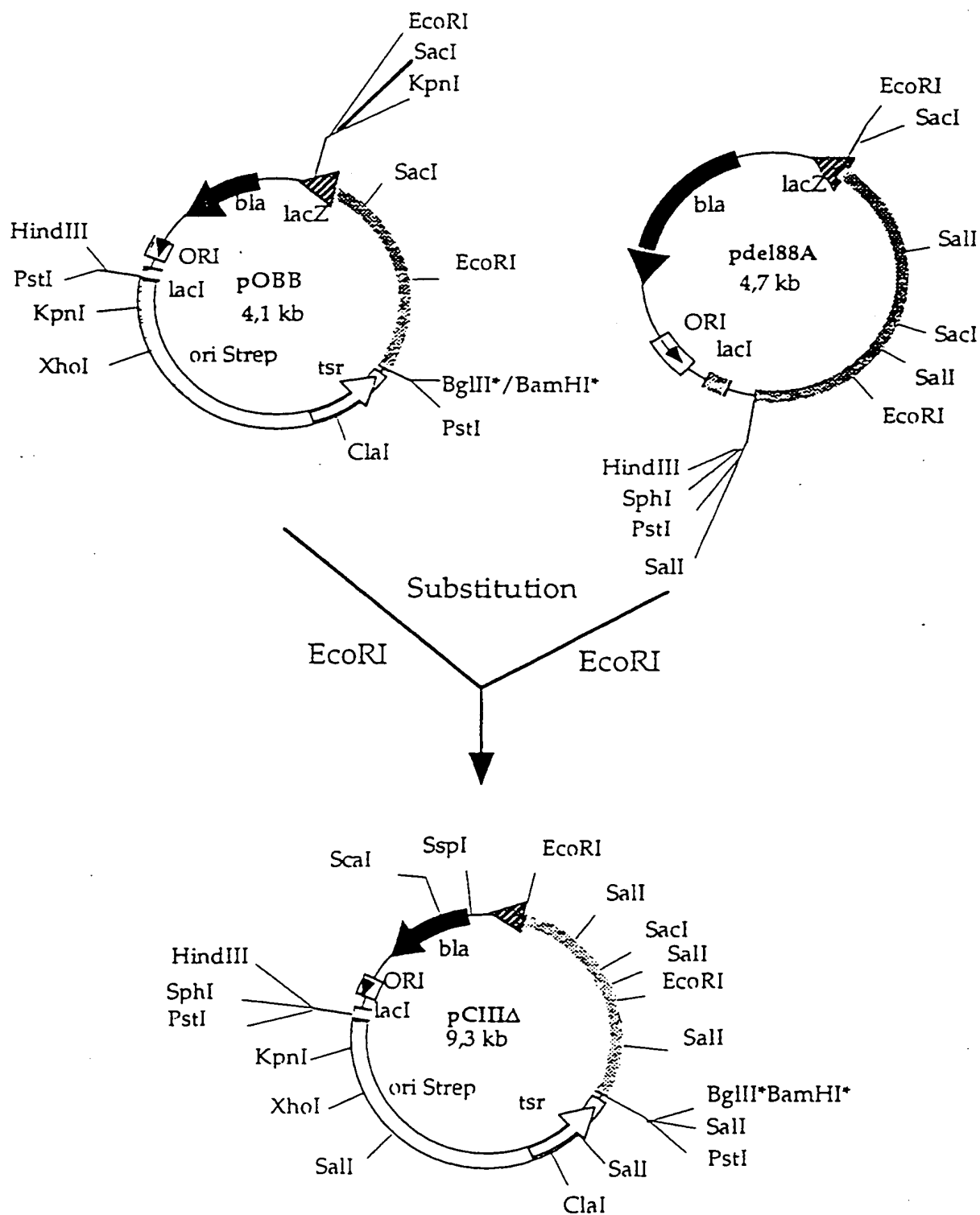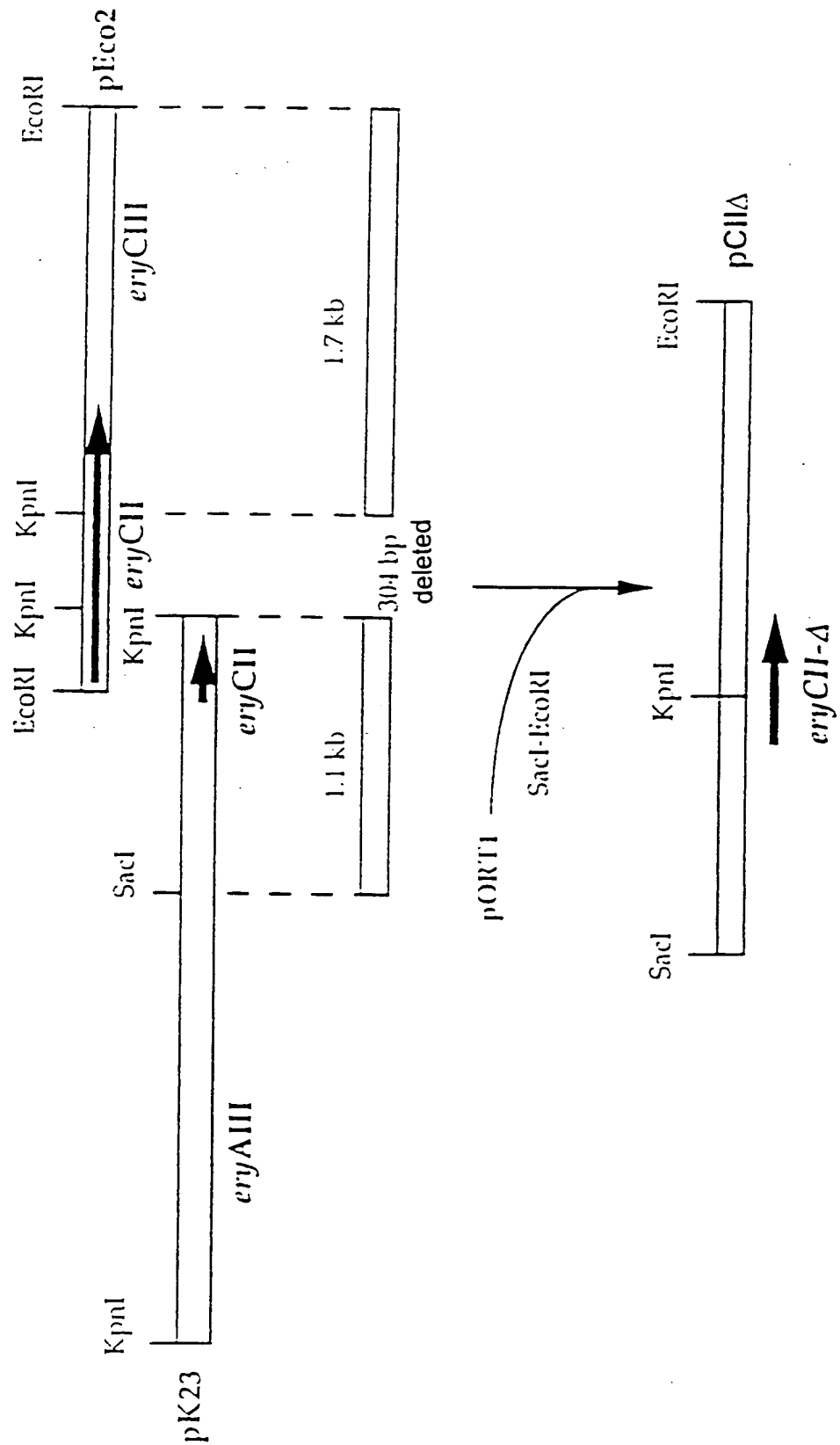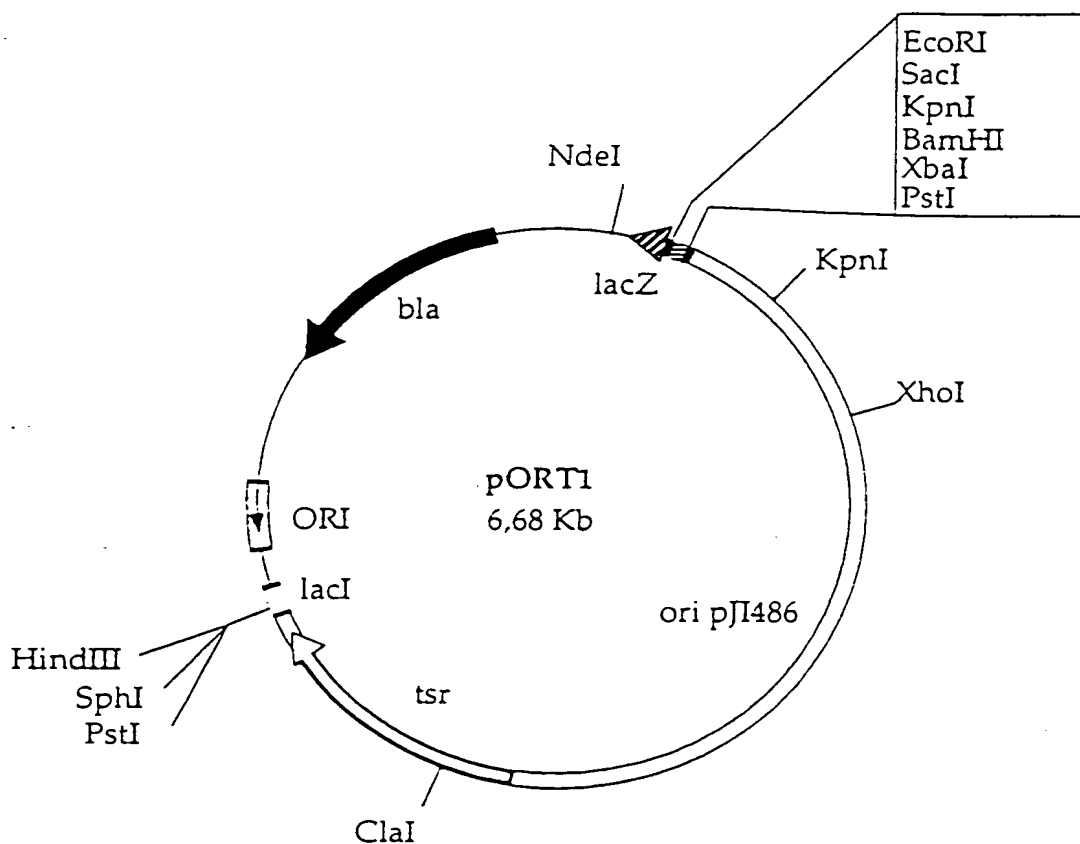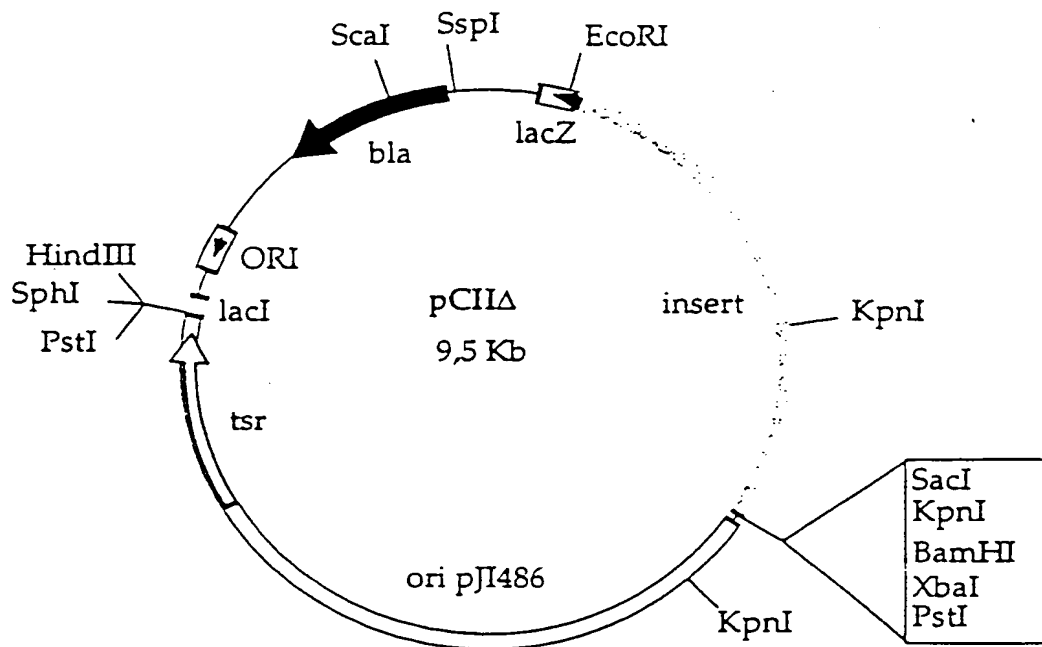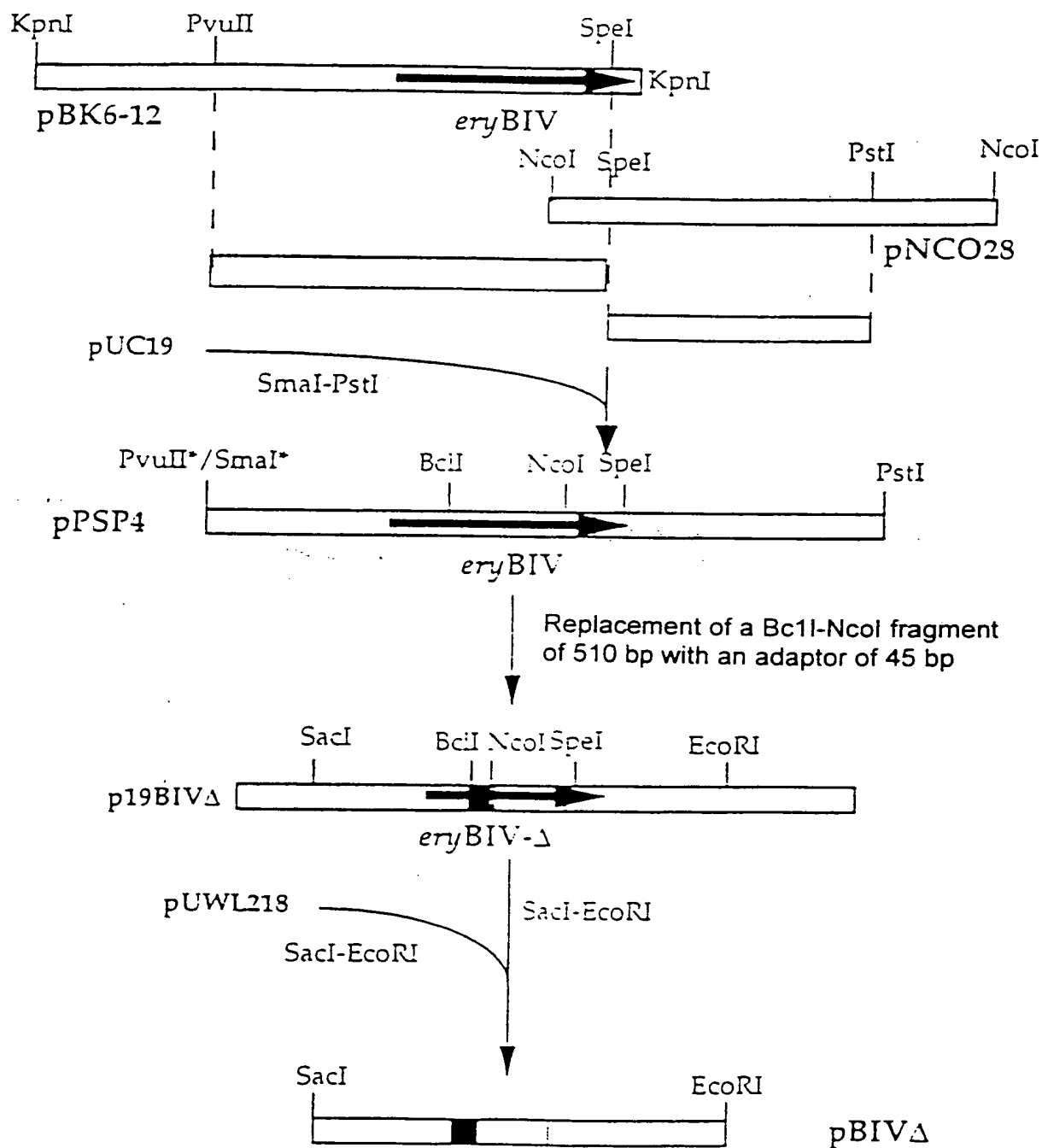
FIGURE 22

721 CGTGACTGCCGGGGCGCTCGCTCCCGGCGCTGGACAGCCTCCTGTGTCCCCAGCAGTTGGCG 780
    R D C R A L A P A L D S L L C P Q Q L A

781 CTGAGCAAGGACATGGCGGTCCGCGCCCTGGAGGACCTGCGTCTCCTCTTCGACGGCCTCGAC 840
    L S K D M A S A L E D L R L L F D G L D

841 GGCGACGCCCGGCCTCGCGGCCCGGCCGACGGTGACGGAACGGCCGTGGCCATGCTCACC 900
    A T P R L A G P A D G D G T A V A M L T

901 GTTCTGCTCTGCACGGAGCCGGTGACCGGCGATGGGAACACCGTGCTCGGGCTCCTT 960
    V L L C T E P V T T A I G N T V L G L L

961 CCCGGGCAGTGGCCGGTGCCCTGCACCGGCCGGGTGGCCGGGCAGGTTGCCGGGGCAG 1020
    P G Q W P V P C T G R V A A G Q V A G Q

1021 GCGCTGCACCGGGGCGGTGTCGTACCGTATCGCCGACGCGGTTCGCGGGAGGACCTGGAG 1080
     A L H R A V S Y R I A T R F A R E D L E

1081 TTGGGGGGCTGGAGGTCAAGTCCGGTGACGAGGTGGTCCTGGCCGAGCGATCGGC 1140
     L A G C E V K S G D E V V V L A G A I G

1141 CGGAACGGACCGTCCGCAGCGCCCCGCCCTCCACCGGGCCCCAGCGGGGCCCCGGCCC 1200
     R N G P S A A P P A P P G P A A P P A

1201 CCGTCGGTCTTCGGTGCCGCCTTCGAGAACGCGGCCCTGGCCGAACCCCTCGTCGGGGCT 1260
     P S V F G A A F E N A L A E P L V R A

1261 GTGACGGGAGCGGCCCTCCAGGCCCTGACGGGCCCTGACCGGGGCGGGGA 1320
     V T G A A L Q A L A E G P P R L T A A G

1321 CCCGTCGTACGACGCGGGCGTTCCCCGTGCGTCGGGGGCTGCCACCGGGCTCCGGTGGCC 1380
     P V V R R R R S P V V G G L H R A P V A

FIGURE 22

```
1381 GCCGCATGAGCATCCGGTCGAACGGGGGCTCGGCGCTCGGGCCCCCTGCGCGGTGA 1440
       A   A   *                                             V   M
                                                              OleG1

1441 TGATGACCACCTTCGCGGCCAACACGCACTTCCAGCCGCTGGTTCCCCTGGCCTGGGCCAC 1500
       M   T   T   F   A   A   N   T   H   F   Q   P   L   V   P   L   A   W   A   L

1501 TGCGGACAGCCGGGCACGAGGTGCGCGTGGTGAGCCAGCCCTCGCTGAGCGACGTGGTGA 1560
       R   T   A   G   H   E   V   R   V   V   S   Q   P   S   L   S   D   V   V   T

1561 CGCAGGCGGGGCTCACCTCGGTCCCCGGTGGGCACCGAGGCTCCGGTCGAGCAGTTCGCGG 1620
       Q   A   G   L   T   S   V   P   V   G   T   E   A   P   V   E   Q   F   A   A

1621 CGACCTGGGGCGGACGATGCCTACATCGGCGTCAACAGCATCGACTTCACCGGCAACGACC 1680
       T   W   G   D   D   A   Y   I   G   V   N   S   I   D   F   T   G   N   D   P

1681 CCGGCCTGTGGACGTGGCCGTACCTCCTGGGCATGGAGACCATGCTGGTGCCGGCCTTCT 1740
       G   L   W   T   W   P   Y   L   L   G   M   E   T   M   L   V   P   A   F   Y

1741 ACGAGTTGCTGAACAACGAGTCCTTCGTGGACGGCGTAGTCGAGTTCGCCCGTGACTGGC 1800
       E   L   L   N   N   E   S   F   V   D   G   V   V   E   F   A   R   D   W   R

1801 GGCCCGACCTGGTGATCTGGGAGCCGCTGACGTTCGCGGCTGGCGCGGCCGGCGTCA 1860
       P   D   L   V   I   W   E   P   L   T   F   A   G   A   V   A   A   R   V   T

1861 CCGGCGGCGGCACGCCCGCCGGTGGGGGCCAGGAGATCACCCTGCGGGGCCGGCAGG 1920
       G   A   H   A   R   L   P   W   G   Q   E   I   T   L   R   G   R   Q   A
                                                                          BamHI

1921 CGTTCCTCCGCGAGCGTGCCCTGCAACCGTTCGAGCACCGGGATCCCACGGCCGAGT 1980
       F   L   A   E   R   A   L   Q   P   F   E   H   R   E   D   P   T   A   E   W

1981 GGCTGGGCCGCATGCTCGACCGGTACGGCTGCTCGTTCGACGAGGAGATGGTCACCGGGC 2040
       L   G   R   M   L   D   R   Y   G   C   S   F   D   E   E   M   V   T   G   Q
```

FIGURE 22

```
2041 AGTGGACCATCGACACGCTGCCCCGCAGCATGCGGCTGGAGCTGTCCGAGGAGCTGCGCA 2100
      W  T  I  D  T  L  P  R  S  M  R  L  E  L  S  E  E  L  R  T

2101 CCCTGGACATGCGTACGTGCCGTACAACGGACCGGGCGGTCGTACCCCCCTGGGTGTGGG 2160
      L  D  M  R  Y  V  P  Y  N  G  P  A  V  V  P  P  W  W  V  W  E

2161 AACCGTGCGAGCGGCGCCCCGGGGTCTGTCTGTGACGATCGGCACCTCCAGCGTGACTCCGGCC 2220
      P  C  E  R  P  R  V  C  L  T  I  G  T  S  Q  R  D  S  G  R

2221 GGGACCATGTCCCCCTCGACCACCTGCTCGACTCCCTGCGGACGCGGGAGATCG 2280
      D  H  V  P  L  D  H  L  L  D  S  L  A  D  V  D  A  E  I  V

2281 TGGCCACGCTCGACACCACCCAGGAGCCCCGGGCCGGGAAGCGCCCCCCGGCAACGTCC 2340
      A  T  L  D  T  T  Q  E  R  L  R  G  A  A  P  G  N  V  R

2341 GGCTGGTGGACTTCGTCCCGCTGCACGCGGCTGATGCCGACCTGCTCGGCGATCGTGCACC 2400
      L  V  D  F  V  P  L  H  A  L  M  P  T  C  S  A  I  V  H  H

2401 ACGGTGGTCCGGGCACGTGGTCGACGGGGCGCTCCACGCGGTCCCCAGATCATCCTGG 2460
      G  G  P  G  T  W  S  T  A  A  L  H  G  V  P  Q  I  I  L  D

2461 ACACCTCGTGGGACACACCGGTGCGGGGCGTGCCAGCGCCATGCAGCAACTCGGGGCGGGGCCTGT 2520
      T  S  W  D  T  P  V  R  A  Q  R  M  Q  Q  L  G  A  G  L  S

2521 CGATGCCGGTGGGGGGAACTGGGCGTCGAGGGCGTCGAGGCCGTCGCTGCGGCTGCTGG 2580
      M  P  V  G  E  L  G  V  E  A  L  R  D  R  V  L  R  L  L  G
                                                    BamHI

2581 GGGAGCCGGAGTTCCGCGCCGGCGGCCGAGCGGCGGCCGATCCGGCCGAGATGCTCGCCATGCCCG 2640
      E  P  E  F  R  A  G  A  E  R  I  R  A  E  M  L  A  M  P  A

2641 CCCCCGGTGACGTCGTACCGGACGACTGGAACGACTCACCGGAGCATGCCACCGGGGCCGA 2700
      P  G  D  V  V  P  D  L  E  R  L  T  A  E  H  A  T  G  A  M
```

FIGURE 22

2701 TGGCGGGAAGGCCGGTGAGACGATGCGCGCTACTGCTGACCTGCTTCGCCAACGACACCCAC 2760
     A G R R *   M R V L L T C F A N D T H
                      OleG2

2761 TTCCACGGGCTGGTGCCCGTGGGGCTGGCGTGGGCGCTGCGGGCACGAAGTCCGCGTG 2820
     F H G L V P L A W A L R A A G H E V R V

2821 GCCAGTCAGCCCGCCCGCCCTGTCCGACACGATCACCCAAGCGGGACTGACCGCGGTGCCCGTG 2880
     A S Q P A L S D T I T Q A G L T A V P V

2881 GGCCCGGGACACCCCGGCCCTTCCTGGAGCTGATGGGGGAGATCGGCGACGTCCAGAAGTAC 2940
     G R D T A F L E L M G E I G A D V Q K Y

2941 TCCACCGGCATCGACCTGGGCGTCCGCGGAGCTGACGAGCTGGGAGTACCTGCTCGGC 3000
     S T G I D L G V R A E L T S W E Y L L G

3001 ATGCACACGACCCTGGTGCCCACGTTCTACTCGCTGGTCAACGACGAGCCGTTCGTCGAC 3060
     M H T T L V P T F Y S L V N D E P F V D

3061 GGGCTCGTCGCGCTGACCCGGGCCTGGCGCCCGGACCTCATCCTGTGGGAGCACTTCAGC 3120
     G L V A L T R A W R P D L I L W E H F S

3121 TTCGCCGGGGCGGTTGGCGGCGGCGCTGGCGGCCCCGCGGTGCTGTGGGGG 3180
     F A G A L A A R A T G T P H A R V L W G

3181 TCGGACCTCATCGTCCGGTTCCGCCGGGACTTCCTCGCGGAACGGGCCGCC 3240
     S D L I V R F R R D F L A E R A N R P A

3241 GAGCACCGGCGAGGACCCCATGGCGGAGTGGCTCGGCTGGGCGGCCGAACGGCTGGGCTCC 3300
     E H R E D P M A E W L G W A A E R L G S

3301 ACCTTCGACGAGGAGCTGGTGACCGGGCAGTGGACGATCGACCCGCTGCCCGGGAGCATG 3360
     T F D E E L V T G Q W T I D P L P R S M

FIGURE 22

```
3361 CGGCTGCCCACCGGGGACGACGACGGTGCCGATGCGGTACGTGCCGTACAACGGGGGGGCC 3420
        R  L  P  T  G  T  T  V  P  M  R  Y  V  P  Y  N  G  R  A

3421 GTGGTCCCCGCCATGGGTCCGGCAGCGTGCCGGCGCCCCCGGATCTGCCTGACGCTCCGGT 3480
        V  V  P  A  W  V  R  Q  R  A  R  R  P  R  I  C  L  T  L  G

3481 GTGTCGGCCCCGGCAGACCCTGGGCGACGGCGTGTCCCTGGCGGAGGTGCTGGCCGCGCTG 3540
        V  S  A  R  Q  T  L  G  D  G  V  S  L  A  E  V  L  A  A  L

3541 GGCGACGTGGACGCCGGAGATCCGTGCCCACGCTGGACGCCTCCCAGGCGAAGCTCCTGGGG 3600
        G  D  V  D  A  E  I  V  A  T  L  D  A  S  Q  R  K  L  L  G

3601 CCGGTGCCGGACAACGTCCCGGCTGGTGGACTTCGTGCCCTGCACGCCCTGATGCCGACC 3660
        P  V  P  D  N  V  R  L  V  D  F  V  P  L  H  A  L  M  P  T

                                    KpnI
3661 TGTTCGGCCGATCGTGCACCAGGGGGCCGGGTACCTGGCTGACGGCCGTCCACGGC 3720
        C  S  A  I  V  H  H  G  G  A  G  T  W  L  T  A  A  V  H  G

3721 GTCCCGCAGATCGTCCTCGGTGACCTCTGGGACAACCTGCTGCGCGCCAGACACAG 3780
        V  P  Q  I  V  L  G  D  L  W  D  N  L  L  R  A  R  Q  T  Q

3781 GCCGCGGGCGGGCCCTGTTCATCCATCCGTCCGAGGTCACCGGGCTCGGTGAG 3840
        A  A  G  A  G  L  F  I  H  P  S  E  V  T  A  A  G  L  G  E

3841 GGCGTGCGCGGGTGCTGACGGACCCTTCCATCCGGGCCGCCACAGCGCGTCCCGGGAC 3900
        G  V  R  R  V  L  T  D  P  S  I  R  A  A  A  Q  R  V  R  D

3901 GAGATGAATGCAGAGCCGACGCCGGAGGTCGTCACGGTGCTGGAGCGGCTCGCCGCCG 3960
        E  M  N  A  E  P  T  P  G  E  V  V  T  V  L  E  R  L  A  A
```

FIGURE 22

09/463705

```
3961  AGCGGCGGGACCCGGACGAGGAGGGGGCGGCGGCGGGAACCATGCGGGGCTGACACGGAGCCGACCACCGG  4020
      S  G  G  R  G  R  G  G  G  N  H  A  G  *
                                    M  R  A  D  T  E  P  T  T  G
                                    oleM

                                    BglII
4021  GTACGAGGACGAGTTCGCCGAGATCTACGACGCCGTGTACCGGGGCCGGGGCAAGGACTA  4080
      Y  E  D  E  F  A  E  I  Y  D  A  V  Y  R  G  R  G  K  D  Y

4081  CGCCGGCGAGGCGAAGGACGTGGCCGACCTCGTGCGCGACCGGGTGCCGGACGCGTCCTC  4140
      A  G  E  A  K  D  V  A  D  L  V  R  D  R  V  P  D  A  S  S

4141  CCTCCTGGACGTGGCCTGCGGCACGGGGGCCGCACTTCGCCACGCTCTTCGA  4200
      L  L  D  V  A  C  G  T  G  A  H  L  R  H  F  A  T  L  F  D

4201  CGACGCCCGGGTCTCGAACTGTCCCGGAGCATGCCCTCCCGCTCCCGCATGCC  4260
      D  A  R  G  L  E  L  S  A  S  M  L  D  I  A  R  S  R  M  P

4261  GGGCGGTGCCGCTGCACCAAGGGACACATGCGATCCTTCGACCTGGGCCCACGGGTCTCCGC  4320
      G  V  P  L  H  Q  G  D  M  R  S  F  D  L  G  P  R  V  S  A

4321  GGTCACCTGCATGTTCAGCTCCGTCGGCCACCTGGCCACCGCCGAACTCGACGCGCGAC  4380
      V  T  C  M  F  S  S  V  G  H  L  A  T  T  A  E  L  D  A  T

4381  GCTGCGGTGCTTCGCCGCCCGGCACACCCGGCCTGGCCGTCATCGAACCGTGGTG  4440
      L  R  C  F  A  R  H  T  R  P  G  G  V  A  V  I  E  P  W  W

4441  GTTCCCCGAGACCTTCACCGACGGGGGTGACATCGTACGGCGTCGACGGCCG  4500
      F  P  E  T  F  T  D  G  Y  V  A  G  D  I  V  R  V  D  G  R

4501  GACCATCTCCCGGGTGTCCCACTCGGTACGGGACGGCGGCACCCGCATGGAGATCCA  4560
      T  I  S  R  V  S  H  S  V  R  D  G  G  A  T  R  M  E  I  H
```

FIGURE 22

```
4561  CTACGTGATCGCCGACGCGCCGAGCACGGTCCCGGCCACCGCCGCATCAC  4620
      Y   V   I   A   D   A   E   H   G   P   R   H   L   V   E   H   H   R   I   T

4621  GCTGTTCCCCGCGGCATGCCTACACGGCCGCGTACGAGAAGGCGGGCTACACCGTCGAGTA  4680
      L   F   P   R   H   A   Y   T   A   A   Y   E   K   A   G   Y   T   V   E   Y

4681  CCTCGACGGGGGCCCTCGGGCCGGGGGCTGTTCGTCGGCACCGGACGTGAACCCGCCC  4740
      L   D   G   G   P   S   G   R   G   L   F   V   G   T   R   T   *

4741  GCGCACCGCCCGATCACCCGCTCAACGCCCGTTCACACGGATCACCGGACCACGCGCGAAGG  4800
                                          M   S   Y   D   D   H   A   V   L   E   A   I   L   R   C   A   G
                                          oleY

4801  ACCTTTCACATGTCGTACGACGACCACGGGTGCTGGAAGCGATACTGCGGTGCGCCGGA  4860
      M   S   Y   D   D   H   A   V   L   E   A   I   L   R   C   A   G

4861  GGTGACGAGGCGGCTTCCTGCTGAACACCGTCGAGGAATGGGAGCGCCGAGATCACCGCG  4920
      G   D   E   R   F   L   N   T   V   E   E   W   G   A   A   E   I   T   A

4921  GCGCTCGTGGACGAGTTGCTGTTCCGCTGCGAGATCCCGCAGTGGGGGTGAGGCGTTC  4980
      A   L   V   D   E   L   L   F   R   C   E   I   P   Q   V   G   G   E   A   F

                                                                    PstI
4981  ATCGGCCTGGACGTCCTGCACGGCGCCGACGGATCAGCCATGTGCTGCAGGTGACGGAC  5040
      I   G   L   D   V   L   H   G   A   D   R   I   S   H   V   L   Q   V   T   D

5041  GGCAAGCCCGTCACGTCGGCGGAACTGGGCGGCCGGTACCTGGAGT  5100
      G   K   P   V   T   S   A   E   P   A   G   Q   E   L   G   G   R   T   W   S

5101  TCACGCTCAGCGACCCTCGCGGGAGCTGTTCGGCCCGCCGTCCGGCCACCGGCGGGGG  5160
      S   R   S   A   T   L   L   R   E   L   F   G   P   P   S   G   R   T   A   G

5161  GGCTTCGGCGTCTCCTTCCTGCCGGATCTGCGCGGACCATGGAGGGCGCGGCC  5220
      G   F   G   V   S   F   L   P   D   L   R   G   P   R   T   M   E   G   A   A
```

FIGURE 22

09/463705

```
5221  CTGGCCGCCCCGCGCCACCAACGTGGTGCTGCACGGCGACGACCAACGAGACGCCCCCACTG 5280
      L  A  A  R  A  T  N  V  V  L  H  A  T  T  N  E  T  P  P  L

5281  GACCGGCTGGCCCTGCGCTACGAGTCCGACAAGTGGGGCGGCGTCCACTGGTTCACCGGC 5340
      D  R  L  A  L  R  Y  E  S  D  K  W  G  G  V  H  W  F  T  G

                                        BamHII
5341  CACTACGACCGGCACCTGCGGGCCGTGCGCGACCAGGCGGTGCGGATCCTGGAGATCGGC 5400
      H  Y  D  R  H  L  R  A  V  R  D  Q  A  V  R  I  L  E  I  G

5401  ATCGGCGGCTACGACGACCTGCTGCCGAGCGGCGCCTCACTGAAGATGTGGAAGCGCTAC 5460
      I  G  G  Y  D  D  L  L  P  S  G  A  S  L  K  M  W  K  R  Y

5461  TTCCCCGCGGCCCTGGTCTTCGGCGTGGACATCTTCGACAGTCGGCGGCTGCGACCAGCCGC 5520
      F  P  R  G  L  V  F  G  V  D  I  F  D  S  R  R  A  T  S  R

5521  GTGTCAAGACGCTCCCGGCCCGGCAGGACCCGGAGTTCATGCGCCGGTCGCCGAG 5580
      V  S  R  R  S  A  A  R  Q  D  D  P  E  F  M  R  R  V  A  E

5581  GAGCACGGCCCGTTCGACGTCATCATCGACGACGGCTCACATCAACGCCACACATGCGG 5640
      E  H  G  P  F  D  V  I  D  D  G  S  H  I  N  A  H  I  M  R

5641  ACGTCGTTCTCGGTGATGTTCCCCCACCTGCGCAACGGGCGCTTCTACGTCATCGAGGAC 5700
      T  S  F  S  V  M  F  P  H  L  R  N  G  G  F  Y  V  I  E  D

5701  ACCTTCACCTCCTACTGGCCCGGGTACGAGGGCCATCCGAGCCCGGTGCCCGTCCGGA 5760
      T  F  T  S  Y  W  P  G  Y  G  G  P  S  G  A  R  C  P  S  G

5761  ACAACCGGCTGGAGATGGTCAAGGGACTGATCGACTCGGTGCACTACGAGGAGGCGGCCG 5820
      T  T  A  L  E  M  V  K  G  L  I  D  S  V  H  Y  E  E  R  P
```

FIGURE 22

```
5821  GACGGGCGCGGCCACGGCCGACTACATCGCCAGGAACCTCGTCGGGCTGCACGGCCTACCAA  5880
         D  G  A  A  T  A  D  Y  I  A  R  N  L  V  G  L  H  A  Y  Q

5881  ACGACCTCGTCTTCCTCGAGAAGGGCGATCAACAAGGAGGGCATCCCCACCACCGTG  5940
         T  T  S  S  S  R  R  A  I  N  K  E  G  G  I  P  H  T  V

5941  CCCCGGGAGCCGTTCTGGAACGACAACTAGCCACGGCCGCAACCAGAGCCGGAAACCGCA  6000
         P  R  E  P  F  W  N  D  N  *

6001  CCACTGTCCCGGCCCACCTCGGAACCACCTCCAGCAAAGGACACACCGCTGTGACCGATAC  6060

                                        KpnI
6061  GCACACCGGACCGACACCGGCCGGGGTACC  6093
```

FIGURE 22